# Sample Paths Estimates for Stochastic Fast-Slow Systems Driven by Fractional Brownian Motion

Katharina Eichinger[1] · Christian Kuehn[1] · Alexandra Neamţu[1]

## Abstract

We analyze the effect of additive fractional noise with Hurst parameter $H > 1/2$ on fast-slow systems. Our strategy is based on sample paths estimates, similar to the approach by Berglund and Gentz in the Brownian motion case. Yet, the setting of fractional Brownian motion does not allow us to use the martingale methods from fast-slow systems with Brownian motion. We thoroughly investigate the case where the deterministic system permits a uniformly hyperbolic stable slow manifold. In this setting, we provide a neighborhood, tailored to the fast-slow structure of the system, that contains the process with high probability. We prove this assertion by providing exponential error estimates on the probability that the system leaves this neighborhood. We also illustrate our results in an example arising in climate modeling, where time-correlated noise processes have become of greater relevance recently.

**Keywords** Fast-slow systems · Fractional Brownian motion · Sample path estimates · Correlated noise · AMOC model

## 1 Introduction

Fast-slow systems naturally arise in the modeling of several phenomena in natural sciences, when processes have widely differing rates [25,27,33]. The standard form of a fast-slow system of ordinary differential equations (ODEs) is given by

$$\frac{\mathrm{d}x}{\mathrm{d}s} = x' = f(x, y, \varepsilon),$$
$$\frac{\mathrm{d}y}{\mathrm{d}s} = y' = \varepsilon g(x, y, \varepsilon), \tag{1}$$

where $x$ are the fast variables, $y$ are the slow variables, $\varepsilon > 0$ is a small parameter, and $f$, $g$ are sufficiently smooth vector fields; for a more detailed technical introduction regarding the analysis of (1) we refer to Sect. 2.1. Here we just point out the basic aspects from the modeling perspective. First, note that if $\varepsilon = 0$, then (1) becomes a parametrized set of ODEs, where the $y$-variables are parameters. Taking this viewpoint, all bifurcation problems [23,35] involving parameters naturally relate to fast-slow dynamics if the parameters vary slowly, which is often a natural assumption in applications. Second, in practice, we also want to couple many dynamical systems. The resulting large/complex system is often multiscale in time and space. For example, in the context of climate modeling [13,28] coupled processes can evolve on temporal scales of seconds up to millennial scales. Third, fast-slow systems are the core class of dynamical problems to understand singular perturbations [54], i.e., roughly speaking singular perturbations problems with small parameters are those, which degenerate in the limit of the small parameter into a different class of equations. Combining all these observations, it is not surprising that fast-slow systems have become an important tool in more theoretical as well as application-oriented parts of nonlinear dynamics [33].

However, when dealing with real life phenomena certain random influences have to be taken into account and quantified in a suitable way [18]. The most common stochastic process used to describe uncertainty is Brownian motion $W = W_t$. One of its key features is the memory-less or Markov property, which means that the behavior of this process after a certain time $T > 0$ only depends on the situation at the current time $T$. In certain applications it may be desirable to model long-range dependencies and to take into account the evolution of the process up to time $T$. One of the most famous example is constituted by fractional Brownian motion (fBm) $W^H = W_t^H$; see [30] for its first use. A fBm is a centered stationary Gaussian processes parameterized by the so-called Hurst index/parameter $H \in (0, 1)$. For $H = 1/2$ one recovers classical Brownian motion. However, for $H \in (1/2, 1)$ and $H \in (0, 1/2)$, fBm exhibits a totally different behavior compared to Brownian motion. Its increments are no longer independent, but positively correlated for $H > 1/2$ and negative correlated for $H < 1/2$. The Hurst index does not only influence the structure of the covariance but also the regularity of the trajectories. Fractional Brownian motion has been used to model a wide range of phenomena such as network traffic [52], stock prices and financial markets [37,50], activity of neurons [17,45], dynamics of the nerve growth [41], fluid dynamics [55], as well as various phenomena in geoscience [31,38,44]. However, the mathematical analysis of stochastic systems involving fBm is a very challenging task. Several well-known results for classical Brownian motion are not available. For instance, the distribution of the hitting time $\tau_a$ of a level $a$ is explicitly known for a Brownian motion, whereas for fBm, one has only an asymptotic statement, according to which

$$\mathbb{P}(\tau_a > t) = t^{-(1-H)+o(1)},$$

as $t$ goes to infinity, see [39]. Furthermore, since fBm is not a semi-martingale, Itô-calculus breaks down. Therefore, it is highly non-trivial to define an appropriate integral with respect to the fBm. This issue has been intensively investigated in the literature. There are numerous approaches that exploit the regularity of the trajectories of the fBm in order to develop a completely path-wise integration theory and to analyze differential equations. For more details, see [19,21,22,26,36] and the references specified therein. Furthermore, another ansatz employed to define stochastic integrals with respect to fBm relies on the stochastic calculus of variations (Malliavin calculus) developed in [11]. In summary, fBm is a natural candidate process to aim to improve our understanding of correlated stochastic dynamics.

Our objective here is to combine the study of fast-slow systems and fBm by starting to study stochastic differential equations of the form

$$
\begin{aligned}
\mathrm{d}x &= \frac{1}{\varepsilon} f(x, y, \varepsilon)\mathrm{d}t + \frac{\sigma}{\varepsilon^H}\mathrm{d}W_t^H, \\
\mathrm{d}y &= 1\mathrm{d}t,
\end{aligned}
\tag{2}
$$

where we start with the case of additive noise for the fast variable(s) and assume there is a single regularly slowly-drifting variable $y$. For $H = 1/2$, i.e., for Brownian motion, there is a very detailed theory, how to analyze stochastic fast-slow systems [33]. One particular building block—initially developed by Berglund and Gentz—uses a sample paths viewpoint [4]. This approach has recently been extended to broader classes of spatial stochastic fast-slow systems [20] and it has found many successful applications; see e.g. [3,32,47,51]. Therefore, it is evident that one should also consider the case of correlated noise in the fast-slow setup [24,57].

Our key goal is to derive sample paths estimates for fast-slow systems driven by fBm with Hurst index $H \in (1/2, 1)$. We restrict ourselves to the case of additive noise and establish the theory for the normally hyperbolic stable case. Due to the technical challenges mentioned above, we need to derive sharp estimates for the exit times for processes solving certain equations driven by fBm. Exploring various properties of general Gaussian processes, we propose two variants to obtain optimal sample paths estimates.

Then we are going to apply our theory to a stochastic climate model [16,42] describing the North-Atlantic thermohaline circulation forced by fractional Brownian motion. In fact, it is well-established that just using white noise modelling in climate models can be insufficient. The simple reason is that neglecting spatial and temporal correlations does not represent the statistics of large classes of underlying climate measurement data including temperature time series [14,29], historical climate data [1,2,10], as well as large-scale simulation data [6]. In all these cases, an elegant way to model temporal correlations in climate science is fractional Brownian motion [1,10,29,49,58]. The reasoning to use a time correlated process can also be understood in climate dynamics in various intuitive ways. For example, in a larger-scale climate model, stochastic terms often represent unresolved degrees of freedom or small-scale fluctuations. If we consider the weather as a short-lived smaller scale effect in terms of the global long-term climate, then models for the latter must include noise with (positive) time correlations as weather patterns are positively correlated in time on short scales [8,56]. Similarly, if the noise terms represent external forcing, such as input from another climate subsystem on a macro-scale, then also this input is likely to be correlated in time as there are internal correlations of the long-term behaviour of each larger-scale climate subsystem. In summary, this has motivated us to consider a model from climate dynamics as one possible key application for fast-slow dynamical systems with fractional Brownian motion. As mentioned above, in many other applications, fractional Brownian motion also naturally appears, so our modelling approach via fast-slow systems with with fBm is even more broadly applicable.

This work is structured as follows. In Sect. 2 we introduce basic notions from the theory of fast-slow systems and fractional Brownian motion. Furthermore, we state important estimates for the exit times of Gaussian processes which will be required later on. In Section 3, we generalize the theory of [4] by first deriving an attracting invariant manifold of the variance using the fast-slow structure of the system. Based on this manifold we define a region, where the linearization of the process is contained with high probability. In order to prove such statements, we first derive a suitable nonlocal Lyapunov-type equation for the covariance of the solution of a linear equation driven by fBm, the so-called fractional Ornstein–Uhlenbeck pro-

cess. Thereafter we analyze two variants which entail sharp estimates for the exit times of this process. Furthermore, we consider more complicated dynamics and provide extensions of our results to the non-linear case, more complicated slow dynamics and finally discuss the case of fully coupled dynamics. We apply our theory to a model for the North-Atlantic thermohaline circulation and provide some simulations. Section 4 generalizes the sample paths estimates to higher dimensions in the autonomous linear case. Our strategy is based on diagonalization techniques, which allow us to go back to the one-dimensional case and apply the results developed in Sect. 3. For completeness, we provide an appendix which contains a detailed proof regarding the limit superior of a non-autonomous fractional Ornstein–Uhlenbeck processes. We conclude in Sect. 5 with an outlook of possible continuations of our results.

## 2 Background

### 2.1 Deterministic Fast-Slow Systems

In this section, we will briefly introduce the terminology of fast-slow systems. We restrict ourselves to the most important results tailored to our problem in the upcoming sections. For further details, see [33]. For the definition of the setting, all of the equations are to be understood formally. We will later add regularity assumptions sufficient to deduce important results. These also imply that the formal computation we will have performed before are valid.

**Definition 2.1** A *fast-slow system* is an (ODE) of the form

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}s} x_s = x'_s &= f(x_s, y_s, \varepsilon), \\
\frac{\mathrm{d}}{\mathrm{d}s} y_s = y'_s &= \varepsilon g(x_s, y_s, \varepsilon),
\end{aligned}
\tag{3}
$$

where $x = x_s$, $y = y_s$ are the unknown functions of the *fast time* variable $s$, the vector fields are $f : \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^m$, $g : \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n$, and $\varepsilon > 0$ is a small parameter. The $x$ variables are called the *fast variables*, while $y$ variables are called the *slow variables*. Transforming into another time scale by defining the *slow time* $t = \varepsilon s$ yields the equivalent system

$$
\begin{aligned}
\varepsilon \frac{\mathrm{d}}{\mathrm{d}t} x_t = \dot{x}_t &= f(x_t, y_t, \varepsilon), \\
\frac{\mathrm{d}}{\mathrm{d}t} y_t = \dot{y}_t &= g(x_t, y_t, \varepsilon).
\end{aligned}
\tag{4}
$$

Depending on the situation both formulations in fast and slow time may be of use. In particular, under certain assumptions, considering them for $\varepsilon \to 0$ indicates a lot of information for the underlying dynamics for the case $0 < \varepsilon \ll 1$. The process for $\varepsilon \to 0$ is called the *singular limit*. The singular limit of (3) for $\varepsilon \to 0$

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}s} x_s = x'_s &= f(x_s, y_s, 0), \\
\frac{\mathrm{d}}{\mathrm{d}s} y_s = y'_s &= 0,
\end{aligned}
$$

is called the *fast subsystem*. The resulting system of the slow time formulation of the fast-slow system (4) for $\varepsilon \to 0$

$$0 = f(x_t, y_t, 0),$$

$$\frac{\mathrm{d}}{\mathrm{d}t} y_t = \dot{y}_t = g(x_t, y_t, 0).$$

is called the *slow subsystem*. The set

$$C_0 := \left\{ (x, y) \in \mathbb{R}^m \times \mathbb{R}^n : \ f(x, y, 0) = 0 \right\}$$

is called the *critical set*. If $C_0$ is a manifold, it is also called the *critical manifold*. From now on, we assume that $C_0$ is a manifold given by a graph of the slow variables, i.e.,

$$C_0 = \left\{ (x^*(y), y) \in \mathbb{R}^m \times \mathbb{R}^n : \ x^* : \mathcal{D} \to \mathbb{R}^m, \ f(x^*(y), y, 0) = 0 \right\},$$

where $\mathcal{D} \subset \mathbb{R}^n$ is an open subset.

**Theorem 2.2** (Fenichel–Tikhonov, [15,27,33,53]) *Let $f, g \in \mathcal{C}^r(\mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R})$, $1 \leq r < \infty$, and their derivatives up to order $r$ be uniformly bounded. Assume that $C_0$ is uniformly hyperbolic. Then for an $\varepsilon_0 > 0$ there exists a locally invariant $\mathcal{C}^r$-smooth manifold*

$$C_\varepsilon = \{ (x, y) : \ x = \bar{x}(y, \varepsilon) \},$$

*for all $\varepsilon \in (0, \varepsilon_0]$, where $\bar{x}(y, \varepsilon) = x^*(y) + \mathcal{O}(\varepsilon)$ with respect to the fast variables. Furthermore, the local stability properties of $C_\varepsilon$ are the same as the ones for $C_0$.*

## 2.2 Fractional Brownian Motion

In this section we state important properties of fBm, which will be required later on. For further details see [5,40] and the references specified therein. We fix a complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and use the abbreviation a.s. for almost surely.

**Definition 2.3** Let $H \in (0, 1]$. A one-dimensional fractional Brownian motion (fBm) of Hurst index/parameter $H$ is a continuous centered Gaussian process $(W_t^H)_{t \geq 0}$ with covariance

$$\mathbb{E}[W_t^H W_s^H] = \frac{1}{2} \left( t^{2H} + s^{2H} - |t - s|^{2H} \right) \qquad \text{for all } t, s \geq 0.$$

Note that for $H > 1/2$ the covariance of fBm satisfies

$$\frac{1}{2}(t^{2H} + s^{2H} - |t - s|^{2H}) = H(2H - 1) \int_0^t \int_0^s |v - u|^{2H-2} \, \mathrm{d}v \, \mathrm{d}u.$$

We further observe that:

(1) for $H = 1/2$ one obtains Brownian motion;
(2) for $H = 1$ then $W_t^H = t W_1^H$ a.s. for all $t \geq 0$. Due to this reason one always considers $H \in (0, 1)$.

The following result regarding the structure of the covariance of fBm holds true, see [40, Section 2.3].

**Proposition 2.4** *Let $H > 1/2$. Then, the covariance of fBm has the integral representation*

$$\mathbb{E}[W_t^H W_s^H] = \int\limits_0^{\min\{s,t\}} K(s,r)K(t,r) \, \mathrm{d}r \quad for \; s, t \geq 0, \tag{5}$$

*where the integral kernel $K$ is given by*

$$K(t,r) = c_H \int\limits_r^t \left(\frac{u}{r}\right)^{H-1/2} (u-r)^{H-3/2} \, \mathrm{d}u,$$

*for a positive constant $c_H$ depending exclusively on the Hurst parameter.*

We remark that for suitable square integrable kernels, one obtains different stochastic processes, for instance the multi-fractional Brownian motion or the Rosenblatt process, see [9]. We now focus on the most important properties of fBm. For the complete proofs of the following statements, see [40, Chapter 2].

**Proposition 2.5** *(Correlation of the increments) Let $(W_t^H)_{t\geq 0}$ be a fBm of Hurst index $H \in (0, 1)$. Then its increments are:*

*(1) positively correlated for $H > 1/2$;*
*(2) independent for $H = 1/2$;*
*(3) negatively correlated for $H < 1/2$.*

*Particularly, for $H > 1/2$ fBm exhibits long-range dependence, i.e.*

$$\sum_{n=1}^{\infty} \mathbb{E}[W_1^H(W_{n+1}^H - W_n^H)] = \infty,$$

*whereas for $H < 1/2$*

$$\sum_{n=1}^{\infty} \mathbb{E}[W_1^H(W_{n+1}^H - W_n^H)] < \infty.$$

**Proposition 2.6** *Let $(W_t^H)_{t\geq 0}$ be a fBm of Hurst index $H \in (0, 1)$. Then:*

*(1) [Self-similarity] For $a \geq 0$*

$$(a^H W_t^H)_{t\geq 0} \stackrel{law}{=} (W_{at}^H)_{t\geq 0}, \tag{6}$$

   *i.e. fBm is self-similar with Hurst index $H$.*

*(2) [Time inversion] $\left(t^{2H} W_{1/t}^H\right)_{t>0} \stackrel{law}{=} (W_t^H)_{t>0}$.*

*(3) [Stationarity of increments] For all $h > 0$*

$$(W_{t+h}^H - W_h^H)_{t\geq 0} \stackrel{law}{=} (W_t^H)_{t\geq 0}.$$

*(4) [Regularity of the increments] fBm has a version which is a.s. Hölder continuous of exponent $\alpha < H$.*

We conclude this section emphasizing the following result, which makes fBm very interesting from the point of view of applications, see [40, Section 2.4 and 2.5].

**Proposition 2.7** *Let $(W_t^H)_{t\geq 0}$ be a fractional Brownian motion with Hurst index $H \in (0, 1/2) \cup (1/2, 1)$. Then $(W_t^H)_{t\geq 0}$ is neither a semi-martingale nor a Markov process.*

### 2.2.1 Integration Theory for $H > 1/2$

Since fBm is not a semi-martingale, the standard Itô calculus is not applicable. Due to this reason, the construction of a stochastic integral of a *random* function with respect to fBm has been a challenging question, see [5,11] and the references specified therein. However, for *deterministic* integrands and for $H > 1/2$ the theory essentially simplifies. We deal exclusively with this case and indicate for the sake of completeness the theory of Wiener integrals of deterministic functions with respect to fBm, see [11]. Let $T > 0$ and

$$\mathcal{E} := \left\{ h : h(s) = \sum_{k=1}^{N-1} h_k \mathbb{1}_{[t_k, t_{k+1})}(s), N \in \mathbb{N}, 0 = t_1 < t_2 < \ldots < t_N \right.$$
$$\left. = T, h_k \in \mathbb{R} \text{ for } k \in \{1, \ldots, N\} \right\}$$

be the set of step functions on $[0, T]$. For $h \in \mathcal{E}$ define the linear mapping $I(h; T) : \mathcal{E} \to L^2(\Omega)$

$$I(h; T) := \int_0^T h(r) \mathrm{d}W_r^H := \sum_{k=1}^{N-1} h_k \left( W_{t_{k+1}}^H - W_{t_k}^H \right).$$

Observe that $I(h; T)$ defines a Gaussian random variable with

$$\mathbb{E}\left[ \int_0^T h(r) \mathrm{d}W_r^H \right] = 0,$$

$$\mathrm{Var}\left[ \int_0^T h(r) \mathrm{d}W_r^H \right] = H(2H-1) \int_0^T \int_0^T h(u)h(v) |u-v|^{2H-2} \, \mathrm{d}u \mathrm{d}v < \infty$$

$$= \int_0^T \int_0^T h(u)h(v)\phi(u-v) \mathrm{d}u \mathrm{d}v, \tag{7}$$

where

$$\phi(s) := H(2H-1) |s|^{2H-2} \tag{8}$$

The representation of the variance can be easily verified by noting the following identity

$$\mathbb{E}\left[ \left( W_{t_{k+1}}^H - W_{t_k}^H \right) \left( W_{t_{l+1}}^H - W_{t_l}^H \right) \right] = H(2H-1) \int_{t_k}^{t_{k+1}} \int_{t_l}^{t_{l+1}} |u-v|^{2H-2} \, \mathrm{d}u \mathrm{d}v.$$

Note that $H > 1/2$ is crucial here. For $p > 1/H$ we can bound the $L^2(\Omega)$-norm of $h \mapsto I(h; T)$ as follows

$$\|I(h; T)\|_{L^2(\Omega)}^2 = H(2H-1) \int_0^T \int_0^T h(u)h(v) |u-v|^{2H-2} \, \mathrm{d}u \mathrm{d}v$$

$$\leq \|h\|_{L^p(0,T)} \|h * \phi\|_{L^{p/(p-1)}(0,T)}$$

$$\leq \|\phi\|_{L^{p/(2p-2)}(0,T)}^2 \|h\|_{L^p(0,T)}^2,$$

where we have obtained the estimate by applying Hölder's inequality and Young's inequality for convolutions [7, Theorem 3.9.4]. The boundedness claim now follows as $\|\phi\|_{L^{p/(2p-2)}(0,T)}^2 < \infty$ for $p > \frac{1}{H}$. This means that $I(\cdot, T)$ is a bounded linear operator defined on the dense subspace $\mathcal{E} \subset L^p(0, T)$, so it can be uniquely extended to a bounded operator

$$I_p(h; T) : L^p(0, T) \to L^2(\Omega).$$

This discussion justifies the following definition:

**Definition 2.8** For $f \in L^p(0, T)$ and $t \in [0, T]$ we set

$$\int_0^t f(r) \mathrm{d}W_r^H := I_p(f \mathbb{1}_{[0,t]}; T)$$

The integral process $\left(I_p(f \mathbb{1}_{[0,t]}; T)\right)_{t \in [0,T]}$ is by construction centered Gaussian. Regarding (7), its covariance can be immediately computed as follows.

**Proposition 2.9** (Covariance of the integral) *Let $a, b > 0$ and $f, g \in L^p(0, T)$ for $p > 1/H$. Then*

$$\mathrm{Cov}\left(\int_0^a f(r) \mathrm{d}W_r^H, \int_0^b g(r) \mathrm{d}W_r^H\right) = \int_0^a \int_0^b f(u)g(v)\phi(u-v) \mathrm{d}u \mathrm{d}v.$$

### 2.2.2 Stochastic Differential Equations Driven by Fractional Brownian Motion

After establishing a suitable stochastic integral with respect to the fractional Brownian motion, we consider stochastic differential equations (SDEs) given by:

$$\mathrm{d}X_t = b(t, X_t)\mathrm{d}t + \sigma(t)\mathrm{d}W_t^H, \quad X_0 = x_0 \in \mathbb{R}, \tag{9}$$

The solution satisfies the integral formulation

$$X_t = x_0 + \int_0^t b(r, X_r)\mathrm{d}r + \int_0^t \sigma(r)\mathrm{d}W_r^H, \text{ a.s.,}$$

where the stochastic integral was constructed in Sect. 2.2.1. Under certain classical regularity assumptions, existence and uniqueness of solutions for (9) can be proven. For more details, see [5, Theorem D.2.4].

**Theorem 2.10** *Let $b : [0, \infty) \times \mathbb{R} \to \mathbb{R}$ be globally Lipschitz in both variables, $\sigma \in \mathcal{C}^1([0, \infty))$ with $\sigma$ and $\frac{\mathrm{d}}{\mathrm{d}t}\sigma$ globally Lipschitz. Then for every $T > 0$ the SDE (9) has a unique continuous solution on $[0, T]$ a.s.*

In this work one case we have to consider is a time-dependent linear drift, i.e., $b(t, \cdot) : \mathbb{R} \to \mathbb{R}$ is linear with $b(t, x) := A(t)x$ for every $t \in [0, \infty)$ and $x \in \mathbb{R}$. In this case, the solution of (9) is given by the variation of constants formula/Duhamel's formula and is called non-autonomous fractional Ornstein–Uhlenbeck process.

**Theorem 2.11** (Non-autonomous Fractional Ornstein–Uhlenbeck Process) *Let $A, B : [0, \infty) \to \mathbb{R}$. Suppose that $A$ is globally Lipschitz and uniformly bounded, and $B \in \mathcal{C}^1([0, \infty))$ with $B$ as well as $\frac{\mathrm{d}}{\mathrm{d}t}B$ globally Lipschitz. Then there exists an a.s. unique solution to the stochastic differential equation*

$$\mathrm{d}X_t = A(t)X_t\mathrm{d}t + B(t)\mathrm{d}W_t^H, \quad X_0 = x_0 \in \mathbb{R} \tag{10}$$

*which satisfies the variation of constants formula*

$$X_t = e^{A(t)}x_0 + \int_0^t e^{\int_r^t A(u)\mathrm{d}u} B(r)\mathrm{d}W_r^H, \text{ a.s.}$$

**Remark 2.12** Note that all the results discussed in this subsection extend to higher dimensions, since all previous steps can be done component-wise. Namely, for $m \geq 1$ we mention.

(R1) We call $(W_t^H)_{t \geq 0}$ an $m$-dimensional fractional Brownian motion if $W_t^H := \sum_{k=1}^{m} W_t^{k,H} e_k$, where $(e_k)_{k \geq 1}$ is a basis in $\mathbb{R}^m$ and $(W_t^{k,H})_{t \geq 0}, k = 1 \ldots m$, are independent one-dimensional fractional Brownian motions with the same Hurst index $H$.

(R2) Naturally, existence and uniqueness of SDEs in higher dimension carry over from Theorem 2.10 under the same assumptions respectively. In particular, for coefficients $A, B : [0, \infty) \to \mathbb{R}^m$ with $m \geq 1$, satisfying the same assumptions as in Theorem 2.11, the solution of (10) is given by

$$X_t = \Phi(t, 0)x_0 + \int_0^t \Phi(t, r)B(r)\mathrm{d}W_r^H, \quad \text{a.s.,}$$

where $\Phi$ denotes the fundamental solution of $x_t' = A(t)x_t$ and $(W_t^H)_{t \geq 0}$ is an $m$-dimensional fractional Brownian motion.

## 2.3 Useful Estimates of Gaussian Processes

The fact that fBm is not a semi-martingale restricts the repository of known inequalities (such as Doob or Burkholder–Davies–Gundy) to establish sample paths estimates. A crucial property of fBm we shall exploit is its Gaussianity. In this section we will describe some useful estimates for exit times of certain Gaussian processes, which will be helpful for our analysis in the upcoming sections.

We first state the next auxiliary result regarding the Laplace transform of a Gaussian process. This was established in [12] by means of Malliavin calculus.

**Lemma 2.13** (Proposition 3.5 [12]) *Let $(Y_t)_{t \geq 0}$ be a centered Gaussian process with $Y_0 = 0$ and covariance function $R(s, t) := \mathbb{E}[Y_s Y_t]$ satisfying the following conditions:*

  (i) $\frac{\partial}{\partial s} R(s, t)$ *exists and is continuous as a function on $[0, \infty) \times [0, \infty)$,*
 (ii) $\frac{\partial}{\partial s} R(s, t) \geq 0$ *for all $t, s \geq 0$,*
(iii) $\mathbb{E}[|Y_t - Y_s|^2] > 0$ *for all $t > s \geq 0$,*
 (iv) $\limsup_{t \to \infty} Y_t = \infty$ *a.s.*

*Then for any $\alpha > 0$:*

$$\mathbb{E}[\exp(-\alpha V_{\tau_c})] \leq \exp(-c\sqrt{2\alpha}), \tag{11}$$

*where $V_t := R(t, t)$ and $\tau_c := \inf\{r > 0 : Y_r \geq c\}$.*

In addition, we require the following form of Chebychev's inequality.

**Lemma 2.14** *Let $\varphi : \mathbb{R} \to [0, \infty)$ be measurable, $Z$ a random variable and $A \in \mathcal{B}(\mathbb{R})$. Then*

$$\inf\{\varphi(y) : y \in A\}\mathbb{P}(Z \in A) \leq \mathbb{E}[\varphi(Z)].$$

*Proof* Under these assumptions we have

$$\inf\{\varphi(y) : y \in A\}\mathbb{1}_{\{Z \in A\}} \leq \varphi(Z)\mathbb{1}_{\{Z \in A\}} \leq \varphi(Z).$$

Taking expectation in the above inequality yields the result. □

**Lemma 2.15** *Let $c > 0$ and $(Y_t)_{t \geq 0}$ be a centered Gaussian process with $Y_0 = 0$ satisfying the assumptions (i)–(iv) of Lemma 2.13. Then, for its exit time $\tau_c := \inf\{r > 0 : Y_r \geq c\}$, the following estimate holds:*

$$\mathbb{P}(\tau_c < t) \leq \exp\left(-\frac{1}{2}\frac{c^2}{\mathrm{Var}(Y_t)}\right).$$

**Proof** Applying Lemma 2.14 for $Z := \tau_c$, $\varphi(r) := \exp(\alpha V_r)$ and $A := (0, t)$ we can bound the probability $\mathbb{P}(\tau_c < t)$ together with (11) as follows:

$$\mathbb{P}(\tau_c < t) \leq \exp\left(\alpha \sup_{0 < r < t} V_r\right) \mathbb{E}\left[\exp\left(-\alpha V_{\tau_c}\right)\right]$$
$$\leq \exp\left(\alpha \sup_{0 < r < t} V_r - c\sqrt{2\alpha}\right),$$

for all $\alpha > 0$. Optimizing over $\alpha$ and noticing that $\sup_{0 < r < t} V_r = \mathrm{Var}(Y_t)$ proves the statement. $\qquad\square$

The previous lemma established a Bernstein-type inequality solely relying on certain properties of the covariance function of Gaussian processes. Another useful estimate is given by [43, Theorem D.4], which is based on Slepian's Lemma [48].

**Theorem 2.16** *Let $T > 0$ and $(Y_t)_{t \in [0,T]}$ be a centered Gaussian process with a.s. continuous trajectories. Assume that $(Y_t)_{t \in [0,T]}$ is a.s. mean-square Hölder continuous, i.e. there are constants $G$ and $\gamma$ such that*

$$\mathbb{E}\left[(Y_t - Y_s)^2\right] \leq G|t - s|^\gamma \quad \text{for all } t, s \in [0, T].$$

*Then there exists a constant $K := K(G, \gamma)$ such that for $c > 0$ and $A \subset [0, T]$*

$$\mathbb{P}\left(\sup_{t \in A} Y_t > c\right) \leq KTc^{\frac{2}{\gamma}} \exp\left(-\frac{c^2}{2\sigma^2(A)}\right),$$

*where $\sigma^2(A) := \sup_{t \in A} \mathrm{Var}(Y_t)$.*

This estimate can be sharpened if we restrict ourselves to the interval of interest.

**Corollary 2.17** *Let $T > 0$ and $(Y_t)_{t \in [0,T]}$ be a centered Gaussian process with a.s. continuous trajectories. Assume that $(Y_t)_{t \in [0,T]}$ is a.s. mean-square Hölder continuous, i.e. there are constants $G$ and $\gamma$, such that*

$$\mathbb{E}\left[(Y_t - Y_s)^2\right] \leq G|t - s|^\gamma \quad \text{for all } t, s \in [0, T].$$

*Then there exists a constant $K := K(G, \gamma)$ such that for $c > 0$ and $0 \leq a < b \leq T$*

$$\mathbb{P}\left(\sup_{a \leq t < b} Y_t > c\right) \leq K(b - a)c^{\frac{2}{\gamma}} \exp\left(-\frac{c^2}{2\sigma^2}\right),$$

*where $\sigma^2 := \sup_{a \leq t < b} \mathrm{Var}(Y_t)$.*

**Proof** $(Z_t)_t$ with $Z_t := Y_{t+a}$ satisfies the assumptions of Theorem 2.16 on $[0, b - a]$. $\qquad\square$

## 3 The One-Dimensional Case

In this section, we investigate the dynamics of a planar stochastic fast-slow system driven by fractional Brownian motion $(W_s^H)_{s \geq 0}$ with Hurst parameter $H > \frac{1}{2}$:

$$
\begin{aligned}
dx_s &= f(x_s, y_s, \varepsilon)ds + \sigma F(y_s)dW_s^H, \\
dy_s &= \varepsilon ds.
\end{aligned}
$$

Its equivalent formulation in slow time, i.e. for $t = \varepsilon s$ is

$$
\begin{aligned}
dx_t &= \frac{1}{\varepsilon} f(x_t, y_t, \varepsilon)dt + \frac{\sigma}{\varepsilon^H} F(y_t)dW_t^H, \\
dy_t &= 1dt,
\end{aligned}
\tag{12}
$$

using the self-similarity of fBm (6). We are interested in the normally hyperbolic stable case and therefore make the following assumptions.

**Assumption 3.1** Stable case

(1) *Regularity:* The functions $f \in C^2(\mathbb{R} \times [0, \infty)^2; \mathbb{R})$ and $F \in C^1([0, \infty); (0, \infty))$, as well as all their existing derivatives up to order two are uniformly bounded on an interval $I = [0, \infty)$ or $I = [0, T]$, $T > 0$, by a constant $M > 0$.
(2) *Critical manifold:* There is an $x^* : [0, \infty) \to \mathbb{R}$ such that

$$
f(x^*(t), t, 0) = 0
$$

for all $t \in [0, \infty)$.
(3) *Stability:* For $a(t) := \partial_x f(x^*(t), t, 0)$ there is $a > 0$ such that

$$
a(t) \leq -a
$$

for all $t \in [0, \infty)$.

Under these assumptions, (12) has a unique global solution according to Theorem 2.10. Furthermore, the deterministic system, i.e., for $\sigma = 0$, given by

$$
\varepsilon \frac{d}{dt} x_t = \varepsilon \dot{x}_t = f(x_t, t, \varepsilon)
$$

has an asymptotically slow manifold $\bar{x}(t, \varepsilon) = x^*(t) + \mathcal{O}(\varepsilon)$ for $\varepsilon > 0$ small enough due to Fenichel–Tikhonov (Theorem 2.2). We expect that, given small noise $0 < \sigma \ll 1$, the trajectories of (12) starting sufficiently close to $\bar{x}(0, \varepsilon)$ remain in a properly chosen neighborhood of $\bar{x}(t, \varepsilon)$ for a long time with high probability. Our goal will be to make this idea rigorous by pursuing the following steps. We first linearize the system around the slow manifold to get an SDE describing the deviations induced by the noise. This helps us obtain a simple description of a suitable neighborhood by using the fast-slow structure inherited by the variance of the system. Then, using this neighborhood, we deduce sample paths estimates for the linear case starting on the slow manifold. To complete the discussion we generalize the result to the non-linear case starting sufficiently close to the slow manifold, that is, such that in the deterministic case solutions are still attracted by the slow manifold. This general strategy inspired by [4], where a similar system driven by Brownian motion (Hurst parameter $H = \frac{1}{2}$) is analyzed. Yet, the several techniques used in [4] do not generalize to fBm.

### 3.1 The Linearized System

The deterministic system

$$\varepsilon \frac{\mathrm{d}}{\mathrm{d}t} x_t = \varepsilon \dot{x}_t = f(x_t, t, \varepsilon)$$

has an asymptotically stable slow manifold $\bar{x}(t, \varepsilon) = x^*(t) + \mathcal{O}(\varepsilon)$ due to Fenichel–Tikhonov (Theorem 2.2). As already outlined, our first step is to examine the behavior of the linearized system around $\bar{x}(t, \varepsilon)$. For a solution $(x_t)_{t \in I}$ of (12) we set $\xi_t := x_t - \bar{x}(t, \varepsilon)$. Then $(\xi)_{t \in I}$ satisfies the equation

$$\begin{aligned}
\mathrm{d}\xi_t &= \frac{1}{\varepsilon} \left[ f(\xi_t + \bar{x}(t, \varepsilon), t, \varepsilon) - f(\bar{x}(t, \varepsilon), t, \varepsilon) \right] \mathrm{d}t + \frac{\sigma}{\varepsilon^H} F(t) \mathrm{d}W_t^H \\
&= \frac{1}{\varepsilon} [a(t, \varepsilon)\xi_t + b(\xi_t, t, \varepsilon)] \mathrm{d}t + \frac{\sigma}{\varepsilon^H} F(t) \mathrm{d}W_t^H,
\end{aligned} \tag{13}$$

where

$$a(t, \varepsilon) = \partial_x f(\bar{x}(t, \varepsilon), t, \varepsilon) = \partial_x f(x^*(t), t, 0) + \mathcal{O}(\varepsilon),$$
$$|b(x, t, \varepsilon)| \leq M |x|^2,$$

by Taylor's remainder theorem. Due to the uniform boundedness of the derivatives of $f$ one can show that the $\mathcal{O}(\varepsilon)$-term is negligible on finite time scales as $\varepsilon \to 0$. Therefore, we restrict ourselves without loss of generality to the analysis of the linearization

$$\mathrm{d}\xi_t = \frac{1}{\varepsilon} a(t)\xi_t \mathrm{d}t + \frac{\sigma}{\varepsilon^H} F(t) \mathrm{d}W_t^H. \tag{14}$$

Examining the process starting on the slow manifold now corresponds to investigating the unique explicit solution of (14) for initial value $\xi_0 = 0$, which is given by the fractional Ornstein–Uhlenbeck process (recall Theorem 2.11)

$$\xi_t = \int_0^t e^{\alpha(t,u)/\varepsilon} \frac{\sigma}{\varepsilon^H} F(u) \mathrm{d}W_u^H,$$

where $\alpha(t, u) := \int_u^t a(r) \mathrm{d}r$. In order to define a proper neighborhood, where the fractional Ornstein–Uhlenbeck process $(\xi_t)_{t \in I}$ is going to stay with high probability, we use the variance $\mathrm{Var}(\xi_t)$ as an indicator for the deviations at time $t$. According to Proposition 2.9, the variance is given by

$$\sigma^2 w(t) := \mathrm{Var}(\xi_t) = \frac{\sigma^2}{\varepsilon^{2H}} \int_0^t \int_0^t e^{\alpha(t,u)/\varepsilon} e^{\alpha(t,v)/\varepsilon} F(u) F(v) H(2H-1) |u-v|^{2H-2} \, \mathrm{d}u \mathrm{d}v.$$

As we would like to see dynamics of $t \mapsto \mathrm{Var}(\xi_t)$, we rescale it by $\frac{1}{\sigma^2}$ to get rid of the small parameter $\sigma \ll 1$, which only changes the order of magnitude of the system. It turns out that $t \mapsto w(t)$ inherits the fast-slow structure from the SDE, which yields a particularly simple approximation of the variance.

**Proposition 3.2** *The so-called renormalized variance $w$ satisfies the fast-slow ODE*

$$\varepsilon \frac{\mathrm{d}}{\mathrm{d}t} w(t) = \varepsilon \dot{w}(t) = 2a(t)w(t) + 2F(t)H(2H-1) \int_0^t \frac{1}{\varepsilon^{2H-1}} e^{\alpha(t,u)/\varepsilon} F(u)(t-u)^{2H-2} \mathrm{d}u. \tag{15}$$

*In particular, there is a (globally) asymptotically stable slow manifold of the system of the form*

$$\zeta(t) = \frac{F(t)^2}{|a(t)|^{2H}} H\Gamma(2H) + \mathcal{O}(\varepsilon). \tag{16}$$

**Proof** Differentiating $t \mapsto w(t)$ yields

$$\frac{d}{dt}w(t) = \dot{w}(t) = 2\frac{a(t)}{\varepsilon}w(t) + 2\frac{F(t)}{\varepsilon^{2H}}H(2H-1)\int_0^t e^{\alpha(t,u)/\varepsilon}F(u)(t-u)^{2H-2}du$$

$$\iff \varepsilon\frac{d}{dt}w(t) = \varepsilon\dot{w}(t) = 2a(t)w(t) + 2F(t)H(2H-1)$$

$$\int_0^t \frac{1}{\varepsilon^{2H-1}}e^{\alpha(t,u)/\varepsilon}F(u)(t-u)^{2H-2}du.$$

In order to be able to take the singular limit $\varepsilon \to 0$ and apply Fenichel–Tikhonov (Theorem 2.2) we need to prove sufficient regularity in $\varepsilon = 0$; continuous differentiability will be enough for the approximation of the slow manifold with the critical manifold up to order $\mathcal{O}(\varepsilon)$. To do this, rewrite the integral by substituting $v = \frac{t-u}{\varepsilon}$

$$\int_0^t \frac{1}{\varepsilon^{2H-1}}e^{\alpha(t,u)/\varepsilon}F(u)(t-u)^{2H-2}du = \int_0^{\frac{t}{\varepsilon}} e^{\alpha(t,t-\varepsilon v)/\varepsilon}F(t-\varepsilon v)v^{2H-2}dv$$

$$\xrightarrow[\varepsilon \to 0]{} F(t)\int_0^\infty e^{a(t)v}v^{(2H-1)-1}dv$$

$$= \frac{F(t)}{|a(t)|^{2H-1}}\Gamma(2H-1).$$

To see that the right hand side of (15) is continuously differentiable in $\varepsilon = 0$ it is sufficient to check it for the integral term

$$\frac{d}{d\varepsilon}\left(\int_0^{\frac{t}{\varepsilon}} e^{\alpha(t,t-\varepsilon v)/\varepsilon}F(t-\varepsilon v)v^{2H-2}dv\right)$$

$$= -\frac{t}{\varepsilon^2}e^{\int_0^{\frac{t}{\varepsilon}} a(t-\varepsilon r)dr}F(0)\left(\frac{t}{\varepsilon}\right)^{2H-2}$$

$$+ \int_0^{\frac{t}{\varepsilon}} e^{\int_0^v a(t-\varepsilon r)dr}\left(-\int_0^v a'(t-\varepsilon r)rdr F(t-\varepsilon v) - F'(t-\varepsilon v)v\right)v^{2H-2}dv,$$

which has an existing limit for $\varepsilon \to 0$, because the exponential term goes to 0 faster than the polynomial term diverges. Now taking the singular limit $\varepsilon \to 0$ gives the slow subsystem

$$0 = 2a(t)w(t) + 2\frac{F(t)^2}{|a(t)|^{2H-1}}H(2H-1)\Gamma(2H-1).$$

The critical manifold is hence given by

$$w^*(t) = \frac{F(t)^2}{|a(t)|^{2H}}H(2H-1)\Gamma(2H-1).$$

Using integration by parts we can rewrite $(2H-1)\Gamma(2H-1) = \Gamma(2H)$, so that the critical manifold can also be written as

$$w^*(t) = \frac{F(t)^2}{|a(t)|^{2H}}H\Gamma(2H).$$

By Theorem 2.2, the ODE (15) has a solution of the form

$$\zeta(t) = \frac{F(t)^2}{|a(t)|^{2H}} H\Gamma(2H) + \mathcal{O}(\varepsilon),$$

which is asymptotically stable due to Assumption 3.13. This stability property is even global in this case because the ODE (15) is linear. □

As expected, the critical manifold depends on the Hurst parameter $H$. Note that as $\Gamma(x) \to 1$ for $x \to 1$, the slow subsystem for $H \to \frac{1}{2}$ reads

$$0 = 2a(t)v(t) + F(t)^2,$$
$$\dot{t} = 1,$$

which coincides with the slow subsystem we would obtain in the case of Brownian motion noise, which exactly corresponds to $H = \frac{1}{2}$.

**Remark 3.3** The proof of Proposition 3.2 only shows that $\zeta$ is $\mathcal{C}^1$ in $\varepsilon$ and $\mathcal{C}^1$ in the time $t$. Depending on the properties of $f$, we expect $\zeta$ to even have higher regularity. However, this fact is not required in the following considerations.

Proposition 3.2 already states that the slow manifold is a good indicator for the size of the set we are looking for as $t \mapsto \frac{1}{\sigma^2}\mathrm{Var}(\xi_t)$ (as a solution of (15) with initial datum $w(0) = 0$) is attracted by the slow manifold. In this particular case we can explicitly state the exponentially fast approach due to the structure of the linear equation

$$\mathrm{Var}(\xi_t) = \sigma^2\left(\zeta(t) - e^{2\alpha(t)/\varepsilon}\zeta(0)\right), \tag{17}$$

where $\alpha(t) := \alpha(t, 0)$. Even more is known about the properties of $\zeta$. Due to the uniform boundedness assumption on $f$ and $F$ we get that the difference between $\zeta$ and $\frac{F(t)^2}{|a(t)|^{2H}} H\Gamma(2H)$ is actually in uniform $t$. This implies that for $\varepsilon$ small enough there are $\zeta^+$ and $\zeta^-$ such that

$$\zeta^+ \geq \zeta(t) \geq \zeta^- > 0 \quad \text{for all } t \in I.$$

The goal is now to prove that the stochastic process $(\xi)_{t \in I}$ is concentrated in sets of the form

$$\mathcal{B}(h) := \{(x, t) \in \mathbb{R} \times I : |x| < h\sqrt{\zeta(t)}\}.$$

To get a better understanding of what to expect, note that the probability that $\xi$ leaves $\mathcal{B}(h)$ at time $t$ can be bounded by using the inequality $\mathbb{P}(X > c) \leq \exp\left(-\frac{c^2}{2\mathrm{Var}(X)}\right)$, which holds for any centered Gaussian random variable $X$. This further leads to

$$\mathbb{P}\left(\frac{|\xi_t|}{\sqrt{\zeta(t)}} \geq h\right) \leq \exp\left(-\frac{h^2\zeta(t)}{2\mathrm{Var}(\xi_t)}\right) \leq \exp\left(-\frac{h^2}{2\sigma^2}\right). \tag{18}$$

Of course, the probability that $(\xi)_{t \in I}$ has exited $\mathcal{B}(h)$ in the interval $[0, t]$ at least once

$$\mathbb{P}\left(\sup_{0 \leq r \leq t} \frac{|\xi_r|}{\sqrt{\zeta(r)}} \geq h\right) = \mathbb{P}\left(\tau_{\mathcal{B}(h)} < t\right)$$

is larger, where $\tau_{\mathcal{B}(h)} := \inf\{r > 0 : (\xi_r, r) \notin \mathcal{B}(h)\}$ is the first time $(\xi)_{t \in I}$ has exited $\mathcal{B}(h)$. We will present a few approaches to estimate this probability in the following, using the inequalities we have established in Section 2.3. The increase of probability compared to (18) is simply indicated by the prefactor of $\exp\left(-\frac{h^2}{2\sigma^2}\right)$.

**Remark 3.4** It could also have been possible to define the neighborhood $\mathcal{B}(h)$ by considering the critical manifold $w^*$, i.e. to define

$$\mathcal{B}^*(h) := \{(x, t) \in \mathbb{R} \times I : |x| < h\sqrt{w^*(t)}\}.$$

This will yield the same bounds on the exit times, which we will establish in the following because the difference between $\zeta(t)$ and $w^*(t)$ is only in $\mathcal{O}(\varepsilon)$, which is of the same order as the order we obtain by approximating with the slow manifold $\zeta$ anyways.

### 3.1.1 Variant 1

The first approach is based on the result on exit times of Gaussian processes with sufficiently regular and increasing covariance function, as stated in Lemma 2.15.

**Theorem 3.5** *Let $t \in I$. Then under Assumption 3.1 for any $h > 0$ the following estimate holds true for $\varepsilon > 0$ sufficiently small*

$$\mathbb{P}\left(\sup_{0 \leq r < t} \frac{|\xi_r|}{\sqrt{\zeta(r)}} \geq h\right) \leq 2e \left\lceil \frac{|\alpha(t)|}{\varepsilon} \frac{h^2}{\sigma^2}(1 + \mathcal{O}(\varepsilon)) \right\rceil \exp\left(-\frac{h^2}{2\sigma^2}\right).$$

**Proof** In order to apply the estimate given in Lemma 2.15 to our problem observe that $\xi$ may not satisfy all the assumptions. First of all we need well-definedness of the Ornstein–Uhlenbeck process over the whole non-negative real line $[0, \infty)$; this is guaranteed by Assumption 3.1. In addition, we consider the process given by $X_t = e^{-\alpha(t)/\varepsilon}\xi_t$. Note that the event that $\xi_t$ exceeds a certain level $c \geq 0$ corresponds exactly to the event of $X_t$ exceeding $e^{-\alpha(t)/\varepsilon}c$, that is $\{\xi_t \geq c\} = \{X_t \geq e^{-\alpha(t)/\varepsilon}c\}$. Unfortunately this observation does not carry over to $\sup_{0 \leq r \leq t} X_r$ and $\sup_{0 \leq r \leq t} \xi_r$. In fact, a priori we only have the relation $\{\sup_{0 \leq r \leq t} \xi_r \geq c\} \subset \{\sup_{0 \leq r \leq t} X_r \geq c\}$, which yields a too strong estimate in the end as we are increasing the variance by an exponentially increasing factor, while maintaining the same exit level! A way to overcome this is to partition the interval $[0, t]$ to suitable subintervals $[t_i, t_{i+1})$ to obtain the relation $\{\sup_{t_i \leq r \leq t_{i+1}} \xi_r \geq c\} \subset \{\sup_{t_i \leq r \leq t_{i+1}} X_r \geq e^{-\alpha(t_i)/\varepsilon}c\}$. This partition will also turn out to be useful to control the variance with the slow manifold in the exponential of the estimate. The covariance of $(X_t)_{t \in I}$ is given by

$$R(t, r) := \mathbb{E}[X_t X_r] = \frac{\sigma^2}{\varepsilon^{2H}} \int_0^t \int_0^r e^{-\alpha(u)/\varepsilon} e^{-\alpha(v)/\varepsilon} F(u)F(v)H(2H-1)\,|u-v|^{2H-2}\,\mathrm{d}u\mathrm{d}v.$$

By the theorem for differentiation of parameter dependent integrals we deduce that $\frac{\partial}{\partial t}R(t, r)$ exists and is continuous in $[0, \infty) \times [0, \infty)$. Furthermore, $\frac{\partial}{\partial t}R(t, r) \geq 0$ for all $t, r \geq 0$ is an immediate consequence of the fundamental theorem of calculus as the integrand is always greater equal than 0. This already implies assumption (i) und (ii) of Lemma 2.15. For assumption (iii) it suffices to observe that

$$X_{t_1} - X_{t_2} = \int_{t_2}^{t_1} e^{-\alpha(u)/\varepsilon} \frac{\sigma}{\varepsilon^H} F(u)\mathrm{d}W_u^H$$

is a Gaussian process with nonzero variance. Assumption (iv) follows by Corollary A.2. This implies that $(e^{-\alpha(t)/\varepsilon}\xi_t)_{t \in I}$ satisfies a Bernstein-type inequality due to Lemma 2.15

$$\mathbb{P}\left(\sup_{0 \leq r \leq t} e^{-\alpha(r)/\varepsilon}\xi_r > c\right) \leq \exp\left(-\frac{1}{2}\frac{c^2}{e^{-2\alpha(t)/\varepsilon}\mathrm{Var}(\xi_t)}\right). \tag{19}$$

After having established this result, we can proceed as in the proof of Proposition 3.1.5 [4]. For $\gamma \in (0, 1/2)$ let $0 = t_0 < t_1 < \ldots < t_N$ be a partition containing the interval $[0, t]$ such that

$$-\alpha(t_{i+1}, t_i) = \varepsilon\gamma \quad \text{for} \quad 0 \le i < N = \left\lceil \frac{|\alpha(t)|}{\varepsilon\gamma} \right\rceil.$$

(Note that $t_N \ge t$ is possible. But this only increases the estimate on the probability slightly. As we would like to optimize over $\gamma$ in the end, it does not make sense to fix it to obtain $t_N = t$.) For each $i \in \{0, \ldots, N-1\}$ we have

$$\mathbb{P}\left( \sup_{t_i \le r < t_{i+1}} \frac{|\xi_r|}{\sqrt{\zeta(r)}} \ge h \right)$$

$$\le 2\mathbb{P}\left( \sup_{0 \le r < t_{i+1}} e^{-\alpha(r)/\varepsilon} \xi_r \ge h \inf_{t_i \le r < t_{i+1}} e^{-\alpha(r)/\varepsilon} \sqrt{\zeta(r)} \right)$$

$$\overset{(19)}{\le} 2\exp\left( -\frac{1}{2} \frac{h^2 \inf_{t_i \le r < t_{i+1}} e^{-2\alpha(r)/\varepsilon} \zeta(r)}{e^{-2\alpha(t_{i+1})/\varepsilon} \operatorname{Var}(\xi_{t_{i+1}})} \right)$$

$$\overset{(17)}{\le} 2\exp\left( -\frac{h^2}{2\sigma^2} \frac{\inf_{t_i \le r < t_{i+1}} \zeta(r)}{\zeta(t_{i+1})} e^{2\alpha(t_{i+1}, t_i)/\varepsilon} \right)$$

$$\le 2\exp\left( -\frac{h^2}{2\sigma^2} e^{-2\gamma} (1 - \mathcal{O}(\varepsilon\gamma)) \right),$$

where the last inequality follows by $\dot{\zeta}(r) = \mathcal{O}(1)$, which is proven in Lemma 3.6. Now by subadditivity of the probability measure

$$\mathbb{P}\left( \sup_{0 \le r < t} \frac{|\xi_r|}{\sqrt{\zeta(r)}} \ge h \right) \le \sum_{i=0}^{N-1} \mathbb{P}\left( \sup_{t_i \le r < t_{i+1}} \frac{|\xi_r|}{\sqrt{\zeta(r)}} \ge h \right)$$

$$\le 2\left\lceil \frac{|\alpha(t)|}{\varepsilon\gamma} \right\rceil \exp\left( \frac{h^2}{\sigma^2} \gamma(1 + \mathcal{O}(\varepsilon)) \right) \exp\left( -\frac{h^2}{2\sigma^2} \right),$$

where the last inequality is due to $e^{-2\gamma} \ge 1 - 2\gamma$. Due to monotonicity of $\lceil \cdot \rceil$, it suffices to minimize

$$\gamma \mapsto \frac{|\alpha(t)|}{\varepsilon\gamma} \exp\left( \frac{h^2}{\sigma^2} \gamma(1 + \mathcal{O}(\varepsilon)) \right)$$

in order to find the minimal value of this estimate. Optimizing over $\gamma$ hence yields

$$\mathbb{P}\left( \sup_{0 \le r < t} \frac{|\xi_r|}{\sqrt{\zeta(r)}} \ge h \right) \le 2e \left\lceil \frac{|\alpha(t)|}{\varepsilon} \frac{h^2}{\sigma^2} (1 + \mathcal{O}(\varepsilon)) \right\rceil \exp\left( -\frac{h^2}{2\sigma^2} \right),$$

which finishes the proof. $\qquad \square$

**Lemma 3.6** *The slow manifold $\zeta(t)$ satisfies $\dot{\zeta}(t) = \mathcal{O}(1)$.*

***Proof*** Note that

$$w^*(t, \varepsilon) = \frac{F(t)}{|a(t)|} H(2H - 1) \int_0^t \frac{1}{\varepsilon^{2H-1}} e^{\alpha(t,u)/\varepsilon} F(u)(t - u)^{2H-2} du$$

satisfies the corresponding invariance equation for the fast-slow ODE (15) up to error $\mathcal{O}(\varepsilon)$. This implies that $\zeta(t) = w^*(t, \varepsilon) + \mathcal{O}(\varepsilon)$. Plugging this representation of $\zeta$ into the ODE (15) yields directly $\varepsilon\dot{\zeta}(t) = \mathcal{O}(\varepsilon)$. $\qquad\square$

### 3.1.2 Variant 2

The second approach uses the fact that $(\xi_t)_{t \in I}$ is mean-square Hölder continuous. This is also going to enable to control the deviations based upon Theorem 2.16.

**Theorem 3.7** *Let $t \in I$. Then under Assumption 3.1 there is a constant $K = K(t, \varepsilon, \sigma, H) > 0$, such that for any $h > 0$ the following estimate holds true for $\varepsilon > 0$ sufficiently small*

$$\mathbb{P}\left(\sup_{0 \le r < t} \frac{|\xi_r|}{\sqrt{\zeta(r)}} \ge h\right)$$

$$\le Kt \exp\left(\frac{h^2}{2\sigma^2}\mathcal{O}(\varepsilon)\right) h^{\frac{1}{H}} \left(\frac{F_+^2}{a^{2H}}H\Gamma(2H) + \mathcal{O}(\varepsilon)\right)^{\frac{1}{2H}} \exp\left(-\frac{h^2}{2\sigma^2}\right),$$

*where $F_+ := \sup_{r \in [0, \infty)} F(r) < \infty$.*

*Proof* Let $t > 0$. In order to apply Theorem 2.16 we have to prove mean-square Hölder continuity of $(\xi_r)_{r \in [0,t]}$. For $t \ge r_1 > r_2 \ge 0$

$$\mathbb{E}\left[\left(\xi_{r_1} - \xi_{r_2}\right)^2\right]$$

$$= \frac{\sigma^2}{\varepsilon^{2H}} \int_0^{r_1} \int_0^{r_1} \left(e^{\alpha(r_1,u)/\varepsilon} - \mathbb{1}_{\{u \le r_2\}}e^{\alpha(r_2,u)/\varepsilon}\right)\left(e^{\alpha(r_1,v)/\varepsilon} - \mathbb{1}_{\{v \le r_2\}}e^{\alpha(r_2,v)/\varepsilon}\right)$$

$$\cdot F(u)F(v)\phi(u-v)\mathrm{d}u\mathrm{d}v$$

$$= \frac{\sigma^2}{\varepsilon^{2H}} \int_0^{r_2} \int_0^{r_2} \left(e^{\alpha(r_1,u)/\varepsilon} - e^{\alpha(r_2,u)/\varepsilon}\right)\left(e^{\alpha(r_1,v)/\varepsilon} - e^{\alpha(r_2,v)/\varepsilon}\right) F(u)F(v)\phi(u-v)\mathrm{d}u\mathrm{d}v \tag{20}$$

$$+ 2\frac{\sigma^2}{\varepsilon^{2H}} \int_{r_2}^{r_1} \int_0^{r_2} \left(e^{\alpha(r_1,u)/\varepsilon} - e^{\alpha(r_2,u)/\varepsilon}\right) e^{\alpha(r_1,v)/\varepsilon} F(u)F(v)\phi(u-v)\mathrm{d}u\mathrm{d}v \tag{21}$$

$$+ \frac{\sigma^2}{\varepsilon^{2H}} \int_{r_2}^{r_1} \int_{r_2}^{r_1} e^{\alpha(r_1,u)/\varepsilon}e^{\alpha(r_1,v)/\varepsilon} F(u)F(v)\phi(u-v)\mathrm{d}u\mathrm{d}v. \tag{22}$$

Lipschitz continuity of $r \mapsto e^{\alpha(r,u)/\varepsilon}$ for arbitrary $u \ge 0$ (with Lipschitz constant $L = \frac{a_+}{\varepsilon}$, where $a_+ := \sup_{r \in I} |a(r)|$) yields for (20)

$$\frac{\sigma^2}{\varepsilon^{2H}} \int_0^{r_2} \int_0^{r_2} \left(e^{\alpha(r_1,u)/\varepsilon} - e^{\alpha(r_2,u)/\varepsilon}\right)\left(e^{\alpha(r_1,v)/\varepsilon} - e^{\alpha(r_2,v)/\varepsilon}\right) F(u)F(v)\phi(u-v)\mathrm{d}u\mathrm{d}v$$

$$\le \frac{\sigma^2 a_+^2}{\varepsilon^{2H+2}} \int_0^{r_2} \int_0^{r_2} F(u)F(v)\phi(u-v)\mathrm{d}u\mathrm{d}v \, |r_1 - r_2|^2$$

$$\le \frac{\sigma^2 a_+^2}{\varepsilon^{2H+2}} F_+^2 \int_0^t \int_0^t \phi(u-v)\mathrm{d}u\mathrm{d}v \, |r_1 - r_2|^2 \le \frac{\sigma^2 a_+^2}{\varepsilon^{2H+2}} F_+^2 t^2 \, |r_1 - r_2|^{2H}.$$

Similarly we can show for (21)

$$2\frac{\sigma^2}{\varepsilon^{2H}}\int_{r_2}^{r_1}\int_0^{r_2}\left(e^{\alpha(r_1,u)/\varepsilon}-e^{\alpha(r_2,u)/\varepsilon}\right)e^{\alpha(r_1,v)/\varepsilon}F(u)F(v)\phi(u-v)\mathrm{d}u\mathrm{d}v$$

$$\leq 2\frac{\sigma^2}{\varepsilon^{2H}}\int_{r_2}^{r_1}e^{\alpha(r_1,v)/\varepsilon}F(v)\int_0^{r_2}F(u)\phi(u-v)\mathrm{d}u\mathrm{d}v\,|r_1-r_2|$$

$$\leq \frac{\sigma^2}{\varepsilon^{2H}}F_+^2\left((r_1-r_2)^{2H}-(r_1^{2H}-r_2^{2H})\right)|r_1-r_2|\leq \frac{\sigma^2}{\varepsilon^{2H}}F_+^2 t(r_1-r_2)^{2H}.$$

Last but not least (22) can be estimated as follows

$$\frac{\sigma^2}{\varepsilon^{2H}}\int_{r_2}^{r_1}\int_{r_2}^{r_1}e^{\alpha(r_1,u)/\varepsilon}e^{\alpha(r_1,v)/\varepsilon}F(u)F(v)\phi(u-v)\mathrm{d}u\mathrm{d}v$$

$$\leq \frac{\sigma^2}{\varepsilon^{2H}}F_+^2\int_{r_2}^{r_1}\int_{r_2}^{r_1}\phi(u-v)\mathrm{d}u\mathrm{d}v=\frac{\sigma^2}{\varepsilon^{2H}}F_+^2\,|r_1-r_2|^{2H}.$$

By combining the three estimates we obtain that, for a constant $G=G(t,\varepsilon,\sigma,H)>0$, it holds

$$\mathbb{E}\left[\left(\xi_{r_1}-\xi_{r_2}\right)^2\right]\leq G\,|r_1-r_2|^{2H}.$$

Let $0=t_0<t_1<\ldots<t_N=t$ be a partition of the interval $[0,t]$ such that

$$t_{i+1}-t_i=\mathcal{O}(\varepsilon)\quad\text{for}\quad 0\leq i<N.$$

For each $i\in\{0,\ldots,N-1\}$ we have by Corollary 2.17 for $K=K(G,2H)$

$$\mathbb{P}\left(\sup_{t_i\leq r<t_{i+1}}\frac{|\xi_r|}{\sqrt{\zeta(r)}}\geq h\right)$$

$$\leq 2\mathbb{P}\left(\sup_{t_i\leq r<t_{i+1}}\xi_r\geq h\inf_{t_i\leq r<t_{i+1}}\sqrt{\zeta(r)}\right)$$

$$\leq K(t_{i+1}-t_i)\left(h\inf_{t_i\leq r<t_{i+1}}\sqrt{\zeta(r)}\right)^{\frac{1}{H}}\exp\left(-\frac{h^2\inf_{t_i\leq r\leq t_{i+1}}\zeta(r)}{2\sup_{t_i\leq r\leq t_{i+1}}\mathrm{Var}(\xi_r)}\right)$$

$$\overset{(17)}{\leq}K(t_{i+1}-t_i)h^{\frac{1}{H}}\left(\sqrt{\frac{F_+^2}{a^{2H}}H\Gamma(2H)+\mathcal{O}(\varepsilon)}\right)^{\frac{1}{H}}\exp\left(-\frac{h^2\inf_{t_i\leq r\leq t_{i+1}}\zeta(r)}{2\sigma^2\sup_{t_i\leq r\leq t_{i+1}}\zeta(r)}\right)$$

$$\leq K(t_{i+1}-t_i)h^{\frac{1}{H}}\left(\frac{F_+^2}{a^{2H}}H\Gamma(2H)+\mathcal{O}(\varepsilon)\right)^{\frac{1}{2H}}\exp\left(-\frac{h^2}{2\sigma^2}(1-\mathcal{O}(\varepsilon))\right),$$

where the last inequality follows by $\dot{\zeta}(r)=\mathcal{O}(1)$, see Lemma 3.6. This yields by the subadditivity of the probability measure

$$\mathbb{P}\left(\sup_{0 \leq r < t} \frac{|\xi_r|}{\sqrt{\zeta(r)}} \geq h\right) \leq \sum_{i=0}^{N-1} \mathbb{P}\left(\sup_{t_i \leq r < t_{i+1}} \frac{|\xi_r|}{\sqrt{\zeta(r)}} \geq h\right)$$

$$\leq K t h^{\frac{1}{H}} \left(\frac{F_+^2}{a^{2H}} H\Gamma(2H) + \mathcal{O}(\varepsilon)\right)^{\frac{1}{2H}} \exp\left(-\frac{h^2}{2\sigma^2}(1 - \mathcal{O}(\varepsilon))\right)$$

$$\leq K t \exp\left(\frac{h^2}{2\sigma^2}\mathcal{O}(\varepsilon)\right) h^{\frac{1}{H}} \left(\frac{F_+^2}{a^{2H}} H\Gamma(2H) + \mathcal{O}(\varepsilon)\right)^{\frac{1}{2H}}$$

$$\exp\left(-\frac{h^2}{2\sigma^2}\right).$$

Therefore, the proof is finished.                                                                                              □

## 3.2 Comparison of the Two Variants

In this section, we will compare the two variants in view of varying the noise intensity given by $\sigma > 0$ and the time scale parameter $\varepsilon > 0$. In order to better understand what to expect under these variations we first heuristically describe their effect on the underlying SDE

$$d\xi_t = \frac{1}{\varepsilon}a(t)\xi_t dt + \frac{\sigma}{\varepsilon^H}F(t)dW_t^H.$$

We can directly see that a smaller $\sigma$ reduces the intensity of the fraction Brownian motion noise. In particular, as $\sigma$ is decreasing, the probability that $(\xi_r)_{r \in I}$ exits $\mathcal{B}(h)$ on some interval $[0, t]$ should become smaller. For a smaller $\varepsilon$ the attraction towards the slow manifold becomes stronger, however also the noise intensity increases. We expect small deviations if $\frac{\sigma}{\varepsilon^H}$ is sufficiently small.

Suppose we are in the situation of Theorems 3.5 and 3.7. To simplify the comparison the results of both variant 1

$$\mathbb{P}\left(\sup_{0 \leq r < t} \frac{|\xi_r|}{\sqrt{\zeta(r)}} \geq h\right) \leq 2e \left\lceil \frac{|\alpha(t)|}{\varepsilon} \frac{h^2}{\sigma^2}(1 + \mathcal{O}(\varepsilon)) \right\rceil \exp\left(-\frac{h^2}{2\sigma^2}\right),$$

and variant 2

$$\mathbb{P}\left(\sup_{0 \leq r < t} \frac{|\xi_r|}{\sqrt{\zeta(r)}} \geq h\right)$$

$$\leq K t \exp\left(\frac{h^2}{2\sigma^2}\mathcal{O}(\varepsilon)\right) h^{\frac{1}{H}} \left(\frac{F_+^2}{a^{2H}} H\Gamma(2H) + \mathcal{O}(\varepsilon)\right)^{\frac{1}{2H}} \exp\left(-\frac{h^2}{2\sigma^2}\right),$$

where $K = K(t, \varepsilon, \sigma, H) > 0$, are displayed here. Unfortunately, we do not know the dependence of $K$ on the other parameters, so we can only do a qualitative comparison up to some extent. By looking at the proof of Theorem 3.7 we guess that $K$ is increasing in $G$, so that we assume for the forthcoming analysis that $K$ is increasing in $\sigma$, decreasing in $\varepsilon$ and increasing in $t$. In variant 1, we see the same interplay of $\sigma$ and $h$ as already observed in the analysis of (18) because the exponential dominates the linear term in the prefactor, and the same holds true for variant 2 as long as $\varepsilon \ll \frac{h^2}{\sigma^2}$ (for $\exp\left(\frac{h^2}{2\sigma^2}\mathcal{O}(\varepsilon)\right)$ to remain relatively small), which is true in many applications. For $\varepsilon \to 0$ the estimate in variant 1 becomes larger, whereas in variant 2 it does not seem to have a huge effect on the bound. However, the

increase might be hidden in $K$. As the time $t$ increases, it obviously becomes more likely that $\xi_t$ has already exited $\mathcal{B}(h)$ at least once. In variant 1 this increase is displayed linearly in $t$ as $r \mapsto a(r)$ is uniformly bounded and thus $\alpha(t) = \Theta(t)$. Variant 2 shows an increase which is at least linear in $t$ because $K$ might be increasing in $t$ as well. This means that in variant 1 we have to pick $h$ large enough such that $\frac{h^2}{\sigma^2}$ is significantly larger than $\ln\left(\frac{t}{\varepsilon}\right)$. For variant 2 we have to choose $h$ in a suitable way that $\frac{h^2}{\sigma^2}$ is larger than $\frac{h^2}{\sigma^2}\mathcal{O}(\varepsilon) + \frac{1}{H}\ln(h) + \ln(Kt)$. Although we cannot prove it, it seems that variant 1 yields a sharper bound. Last but not least, note that the estimate in variant 1 coincides with the estimate derived for the Brownian motion case, see [4, Proposition 3.1.5]. The dependence of the Hurst parameter $H$ is completely hidden in the structure of the neighborhood $\mathcal{B}(h)$, which depends on the slow manifold. This also intuitively makes sense because we are "almost" dividing by the variance. Furthermore, in the Brownian motion case this estimate is quite close to the actual distribution of the exit time $\tau_{\mathcal{B}(h)}$, see [4, Theorem 3.1.6] and the comments below.

## 3.3 Back to the Original System

Now that we have convinced ourselves that the most promising estimate is given in Theorem 3.5 it remains to generalize the result to different scenarios which may be of interest.

### 3.3.1 The Nonlinear Case

Recall that we have rewritten the SDE (13) satisfied by the deviations around the slow manifold $\xi_t = x_t - \bar{x}(t, \varepsilon)$

$$d\xi_t = \frac{1}{\varepsilon}[a(t, \varepsilon)\xi_t + b(\xi_t, t, \varepsilon)]dt + \frac{\sigma}{\varepsilon^H}F(t)dW_t^H,$$

with $a$ being the linear drift term and $b$ containing the (possible) nonlinearities of the equation satisfying $|b(x, t, \varepsilon)| \leq M |x|^2$. As we expect that the nonlinear term $b$ does not influence the deviations too strongly near the critical manifold, we use the same neighborhood $\mathcal{B}(h)$. This in particular implies that we keep the same simple description of it, which we have derived in Proposition 3.2. The bound on $b$ will help us to control it inside of $\mathcal{B}(h)$. For the case that $(x_t)_{t \in I}$ is starting on the slow manifold, i.e. $\xi_0 = 0$ we obtain:

**Theorem 3.8** *Let $t \in I$. For $h$ sufficiently small it holds*

$$\mathbb{P}\left(\sup_{0 \leq r < t} \frac{|\xi_r|}{\sqrt{\zeta(r)}} \geq h\right) \leq 2e \left\lceil \frac{|\alpha(t)|}{\varepsilon}\kappa^2\frac{h^2}{\sigma^2}(1 + \mathcal{O}(\varepsilon)) \right\rceil \exp\left(-\kappa^2\frac{h^2}{2\sigma^2}\right),$$

*where $\kappa = 1 - \mathcal{O}(h)$.*

***Proof*** As previously motivated before we treat the nonlinear drift term $b$ as perturbation of the linear system, i.e. split the solution of (13)

$$\xi_t = \xi_t^0 + \xi_t^1,$$

where $\xi_t^0$ is a solution to the linear system, which we have already studied in detail, and

$$\xi_t^1 = \frac{1}{\varepsilon}\int_0^t e^{\alpha(t,u)/\varepsilon}b(\xi_u, u, \varepsilon)du.$$

Then for $h = h_0 + h_1$, $h_0, h_1 \geq 0$ we consider

$$\mathbb{P}\left(\sup_{0 \leq r < t} \frac{|\xi_r|}{\sqrt{\zeta(r)}} \geq h\right) = \mathbb{P}\left(\sup_{0 \leq r < t \wedge \tau_{\mathcal{B}(h)}} \frac{|\xi_r|}{\sqrt{\zeta(r)}} \geq h\right)$$

$$\leq \underbrace{\mathbb{P}\left(\sup_{0 \leq r < t} \frac{|\xi_r^0|}{\sqrt{\zeta(r)}} \geq h_0\right)}_{=:P_0(h_0)} + \underbrace{\mathbb{P}\left(\sup_{0 \leq r < t \wedge \tau_{\mathcal{B}(h)}} \frac{|\xi_r^1|}{\sqrt{\zeta(r)}} \geq h_1\right)}_{=:P_1(h_1)}$$

It remains to prove that $P_1(h_1)$ is small. Observe that due to continuity of $(\xi_t)_{t \in I}$ we have for $r < \tau_{\mathcal{B}(h)}$

$$|\xi_r| \leq h\sqrt{\zeta(r)} \leq h\sqrt{\zeta^+}.$$

This enables us to control $\frac{|\xi_r^1|}{\sqrt{\zeta(r)}}$ inside of $\mathcal{B}(h)$ thanks to the bound on the Taylor remainder term $b$

$$\frac{|\xi_r^1|}{\sqrt{\zeta(r)}} \leq \frac{1}{\varepsilon\sqrt{\zeta(r)}} \int_0^r e^{\alpha(r,u)/\varepsilon} |b(\xi_u, u, \varepsilon)| \, \mathrm{d}u$$

$$\leq \frac{M}{\varepsilon\sqrt{\zeta(r)}} \int_0^r e^{\alpha(r,u)/\varepsilon} \underbrace{|\xi_u|^2}_{\leq h^2 \zeta^+} \, \mathrm{d}u \leq \frac{Mh^2}{a} \frac{\zeta^+}{\sqrt{\zeta^-}}.$$

Hence, choosing $h_1 = 2\frac{Mh^2}{a} \frac{\zeta^+}{\sqrt{\zeta^-}}$ results in $P_1(h_1) = 0$. Note that this is choice is possible as long as $h \leq \frac{a}{2M} \frac{\sqrt{\zeta^-}}{\zeta^+}$, which is in $\mathcal{O}(1)$, so requiring $h \gg \sigma$ is possible. Indeed, the choice of $h$ is usually even "smaller" than $\mathcal{O}(1)$, so that $h_0 = h - h_1 = h(1 - \mathcal{O}(h))$. Applying Theorem 3.5 to $h_0$ now yields the claim. □

**Remark 3.9** With this approach we lose some accuracy ($\kappa = 1 - \mathcal{O}(h)$ instead of $\kappa = 1$) in the exponential. This has more effect on the increase of probability than in the prefactor. To overcome this difficulty it might be better to adapt the neighborhood depending on the nonlinearities.

### 3.3.2 Behavior Close to the Slow Manifold

For the deterministic system we get in the case of a uniformly asymptotically stable slow manifold that solutions starting close to it are attracted exponentially fast. Given low enough noise intensity a similar behavior can be observed in the noisy system, i.e., solutions have small deviations around the deterministic solution and after some (small) time $t_0$ we can again observe small deviations around the slow manifold.

**Theorem 3.10** *Let $t > 0$, $d > 0$. There is $\delta > 0$ and some time $t_0 > 0$ such that the solutions $(x_t)_{t \in I}$ of (12) with initial condition $x_0$ satisfying*

$$|x_0 - \bar{x}(0, \varepsilon)| < \delta$$

*are attracted by the slow manifold. That is, up to time $t_0$ the solution $(x_t)_{t \in I}$ is close to the deterministic solution $x^{\mathrm{det}}$*

$$\mathbb{P}\left(\sup_{0\le r<t_0}\frac{|x_r-x_r^{\mathrm{det}}|}{\sqrt{\tilde{\zeta}(r)}}\ge h\right)\le 2e\left\lceil\frac{|\tilde{\alpha}(t_0)|}{\varepsilon}\kappa^2\frac{h^2}{\sigma^2}(1+\mathcal{O}(\varepsilon))\right\rceil\exp\left(-\kappa^2\frac{h^2}{2\sigma^2}\right),$$

where the ˜ denotes the different values due to linearization around $x^{\mathrm{det}}$ instead of the slow manifold $\bar{x}(t,\varepsilon)$. After $t_0$ we obtain almost the same behavior as in the case where $x_0 = x(0,\varepsilon)$, i.e. for $t \ge t_0$

$$\mathbb{P}\left(\sup_{t_0\le r<t}\frac{|x_r-\bar{x}(r,\varepsilon)|}{\sqrt{\zeta(r)}}\ge h\right)\le 2e\left\lceil\frac{|\alpha(t)|}{\varepsilon}\kappa^2\frac{(h-d)^2}{\sigma^2}(1+\mathcal{O}(\varepsilon))\right\rceil\exp\left(-\kappa^2\frac{(h-d)^2}{2\sigma^2}\right).$$

**Proof** Exponentially fast attraction means that there are constants $\delta, C, \kappa > 0$ such that for $x_0 \in \mathbb{R}$ with $|x_0 - \bar{x}(0,\varepsilon)| < \delta$ it holds

$$|x_t - x(t,\varepsilon)| \le C\,|x_0 - \bar{x}(0,\varepsilon)|\,e^{-\kappa t/\varepsilon}.$$

Consider an initial value $x_0 \in \mathbb{R}$ with $|x_0 - x(0,\varepsilon)| < \delta$ and denote by $x^{\mathrm{det}}$ the solution to the deterministic system ($\sigma = 0$) in (12). Instead of linearizing (12) around $\bar{x}(t,\varepsilon)$, like we did in (13), we linearize it around $x^{\mathrm{det}}$. This procedure yields qualitatively the same linearization, with $\tilde{a}(t) := \partial_x f(x_t^{\mathrm{det}}, t, 0)$ instead of $a(t,\varepsilon)$, or respectively $a(t)$, see discussion before, and $\tilde{\zeta}$ adapted accordingly. In particular, even for the nonlinear case we obtain by Theorem 3.8

$$\mathbb{P}\left(\sup_{0\le r<t}\frac{|x_r-x_r^{\mathrm{det}}|}{\sqrt{\tilde{\zeta}(r)}}\ge h\right)\le 2e\left\lceil\frac{|\tilde{\alpha}(t)|}{\varepsilon}\kappa^2\frac{h^2}{\sigma^2}(1+\mathcal{O}(\varepsilon))\right\rceil\exp\left(-\kappa^2\frac{h^2}{2\sigma^2}\right),$$

where $\tilde{\alpha}(t) = \int_0^t \tilde{a}(r)\mathrm{d}r$ and $\kappa = 1 - \mathcal{O}(h)$. Choose $t_0$ such that for distance $d$

$$\left|x_r^{\mathrm{det}} - \bar{x}(r,\varepsilon)\right| < d \ \text{ for all } r \ge t_0,$$

that is $t_0 \ge \frac{\varepsilon}{\kappa}\ln\left(\frac{C\delta}{d}\right)$. Furthermore, we have by the mean value theorem

$$\tilde{a}(t,\varepsilon) = \partial_x f(x_t^{\mathrm{det}}, t, \varepsilon) = \partial_x f(\bar{x}(t,\varepsilon), t, \varepsilon) + \partial_{xx} f(\tilde{x}, t, \varepsilon)(x_t^{\mathrm{det}} - \bar{x}(t,\varepsilon))$$

for some $\tilde{x} = \lambda x_t + (1-\lambda)\bar{x}(t,\varepsilon)$, $\lambda \in [0,1]$. So that for $a(t,\varepsilon) = \partial_x f(\bar{x}(t,\varepsilon), t, \varepsilon)$

$$|\tilde{a}(t,\varepsilon) - a(t,\varepsilon)| \le CM\delta e^{-\kappa t/\varepsilon}.$$

We want this distance to be of order at most $\varepsilon$, so that in total $t_0 \ge \max\left\{\frac{\varepsilon}{\kappa}\ln\left(\frac{C\delta}{d}\right), \frac{\varepsilon}{\kappa}\ln\left(\frac{CM\delta}{\varepsilon}\right)\right\}$. Then, up to time $t_0$, we can use the estimate above for $(x_t)_{t\in I}$ close to its deterministic solution. And after $t_0$ the process is already close to the slow manifold and its dynamics, so it makes sense to look at the deviations around $\bar{x}(r,\varepsilon)$. Splitting again $h = h_0 + h_1$, $h_0, h_1 \ge 0$

$$\mathbb{P}\left(\sup_{t_0\le r<t}\frac{|x_r-\bar{x}(r,\varepsilon)|}{\sqrt{\zeta(r)}}\ge h\right)\le\mathbb{P}\left(\sup_{t_0\le r<t}\frac{|x_r-x_r^{\mathrm{det}}|}{\sqrt{\zeta(r)}}\ge h_0\right)$$
$$+\mathbb{P}\left(\sup_{t_0\le r<t}\frac{|x_r^{\mathrm{det}}-\bar{x}(r,\varepsilon)|}{\sqrt{\zeta(r)}}\ge h_1\right).$$

Choosing $h_1 = d$ we obtain $\mathbb{P}\left(\sup_{t_0 \le r < t} \frac{|\bar{x}(r,\varepsilon) - x_r^{\text{det}}|}{\sqrt{\zeta(r)}} \ge h_1\right) = 0$. Furthermore, since $h_0 = h - d$, we have

$$\mathbb{P}\left(\sup_{t_0 \le r < t} \frac{|x_r - \bar{x}(r,\varepsilon)|}{\sqrt{\zeta(r)}} \ge h\right) \le 2e\left[\frac{|\alpha(t)|}{\varepsilon}\kappa^2\frac{(h-d)^2}{\sigma^2}(1+\mathcal{O}(\varepsilon))\right]\exp\left(-\kappa^2\frac{(h-d)^2}{2\sigma^2}\right),$$

which finishes the proof.                                                                                      □

### 3.3.3 More Complicated Slow Dynamics

So far, we have considered the case where the slow dynamics is completely uniform and regular, i.e.,

$$dy_t = 1dt.$$

However, in applications many interesting systems contain more complicated slow variables. In fact, this is particularly relevant if one wants to reduce the dynamics to the slow manifold. The reduced equation usually qualitatively describes the dynamics of the slow variables around the slow manifold quite well. This section will clarify that the theory developed so far can be extended to more complicated slow dynamics in two steps. We first generalize our result to deterministic slow dynamics, which may also influence the diffusion term and then consider a fully coupled system.

### Deterministic Slow Dynamics

Hence, we consider systems of the form

$$
\begin{aligned}
dx_t &= \frac{1}{\varepsilon}f(x_t, y_t, \varepsilon)dt + \frac{\sigma}{\varepsilon^H}F(y_t)dW_t^H \\
dy_t &= g(y_t, \varepsilon)dt.
\end{aligned}
\tag{23}
$$

We have to adapt the assumptions a bit.

**Assumption 3.11** Stable case, non-trivial slow dynamics

1. *Regularity* The functions $f \in \mathcal{C}^2(\mathcal{D} \times [0, \infty); \mathbb{R})$, $g \in \mathcal{C}^2(\pi_2(\mathcal{D}) \times [0, \infty); \mathbb{R})$ and $F \in \mathcal{C}^1(\pi_2(\mathcal{D}); (0, \infty))$, as well as their derivatives up to order 2 are uniformly bounded on an open subset $\mathcal{D} \subset \mathbb{R}^2$ by a constant $M \ge 0$. Here $\pi_2$ is the projection onto the second coordinate.
2. *Critical manifold* There is an $x^* : \mathcal{D}_0 \to \mathbb{R}$ for $\mathcal{D}_0 \subset \pi_2(\mathcal{D})$ open such that

$$C_0 = \{(x, y) \in \mathcal{D} : \ y \in \mathcal{D}_0, \ x = x^*(y)\}$$

   is a critical manifold of the system (23).
3. *Stability* For $a(y) := \partial_x f(x^*(y), y, 0)$ there is $a > 0$ such that

$$a(y) \le -a$$

   for all $y \in \mathcal{D}_0$.
4. *Global existence* The solutions $(x, y)$ of (23) are defined for all $t \in [0, \infty)$.

Under these assumptions the system (23) has an attracting slow manifold

$$C_\varepsilon = \{(x, y) \in \mathcal{D} : \ y \in \mathcal{D}_0, \ x = \bar{x}(y, \varepsilon)\},$$

where $\bar{x}(y, \varepsilon) = x^*(y) + \mathcal{O}(\varepsilon)$ due to Theorem 2.2 (Fenichel–Tikhonov). (Again, this $\mathcal{O}(\varepsilon)$ is uniform in $y$.) We linearize the fast variable around $C_\varepsilon$. For a solution $(x, y)$ of (23) set $\xi_t = x_t - \bar{x}(y_t, \varepsilon)$, then $(\xi, y)$ satisfies the equation

$$\begin{aligned}
\mathrm{d}\xi_t &= \frac{1}{\varepsilon} \left[ f(\xi_t + \bar{x}(y_t, \varepsilon), y_t, \varepsilon) - f(\bar{x}(y_t, \varepsilon), y_t, \varepsilon) \right] \mathrm{d}t + \frac{\sigma}{\varepsilon^H} F(y_t) \mathrm{d}W_t^H \\
&= \frac{1}{\varepsilon} [a(y_t, \varepsilon)\xi_t + b(\xi_t, y_t, \varepsilon)] \mathrm{d}t + \frac{\sigma}{\varepsilon^H} F(y_t) \mathrm{d}W_t^H, \\
\mathrm{d}y_t &= g(y_t, \varepsilon)\mathrm{d}t,
\end{aligned} \tag{24}$$

where due to Taylor's remainder theorem

$$a(y, \varepsilon) = \partial_x f(\bar{x}(y, \varepsilon), y, \varepsilon) = \partial_x f(x^*(y), y, 0) + \mathcal{O}(\varepsilon),$$
$$|b(x, y, \varepsilon)| \le M \, |x|^2 \,.$$

Now we can proceed as in the case for trivial slow dynamics by first considering solutions starting on the slow manifold, i.e. $(\xi_0, y_0) = (0, y_0), \ y_0 \in \mathcal{D}_0$, and using the terms

$$\tilde{a}(t) := a(y_t, 0),$$
$$\tilde{F}(t) := F(y_t).$$

This way we obtain the same qualitative bound (also for the nonlinear case) as before, which also coincides with intuition, as more involved dynamics on the slow manifold should not influence the attracting behavior of it.

## Fully Coupled Dynamics

Now that we have seen the idea how to generalize to more complicated dynamics we give an exposition of the more general case, where the slow variables may even be random, particularly be perturbed by fBms with different Hurst parameters $H_1, H_2 \in (\frac{1}{2}, 1)$. We consider the following system

$$\begin{aligned}
\mathrm{d}x_t &= \frac{1}{\varepsilon} f(x_t, y_t, \varepsilon)\mathrm{d}t + \frac{\sigma_1}{\varepsilon^{H_1}} \mathrm{d}W_t^{H_1} \\
\mathrm{d}y_t &= g(x_t, y_t, \varepsilon)\mathrm{d}t + \sigma_2 \mathrm{d}W_t^{H_2},
\end{aligned} \tag{25}$$

which will turn out to be interesting in applications, see Sect. 3.4. The following assumptions will suffice to obtain a qualitatively similar result to the one-dimensional case, analyzed in Sect. 3.

**Assumption 3.12** Stable case, fully coupled system

1. *Regularity* The functions $f \in \mathcal{C}^2(\mathcal{D} \times [0, \infty); \mathbb{R})$ and $g \in \mathcal{C}^2(\mathcal{D} \times [0, \infty); \mathbb{R})$ as well as their derivatives up to order 2 are uniformly bounded on an open subset $\mathcal{D} \subset \mathbb{R}^2$.
2. *Critical manifold* There is an $x^* : \mathcal{D}_0 \to \mathbb{R}$ for $\mathcal{D}_0 \subset \pi_2(\mathcal{D})$ open such that

$$C_0 = \{(x, y) \in \mathcal{D} : \ y \in \mathcal{D}_0, \ x = x^*(y)\}$$

is a critical manifold of the system (25). Here $\pi_2$ is the projection onto the second coordinate.

3. *Stability* For $a(y) := \partial_x f(x^*(y), y, 0)$ there is $a > 0$ such that

$$a(y) \le -a$$

for all $y \in \mathcal{D}_0$.
4. *Global existence* The solutions $(x, y)$ of (25) are defined for all $t \in [0, \infty)$.

**Remark 3.13** Note that the theory discussed in Sect. 2.2.1 does not provide the technical details regarding the existence and uniqueness of solutions for coupled systems driven by fractional Brownian motion. However, this can be extended to systems of the form (25). We refer to [40, Theorem 3.3] for further details.

Similarly as before the system (25) has an attracting slow manifold given by

$$C_\varepsilon = \{(x, y) \in \mathcal{D} : y \in \mathcal{D}_0, \ x = \bar{x}(y, \varepsilon)\},$$

where $\bar{x}(y, \varepsilon) = x^*(y) + \mathcal{O}(\varepsilon)$ due to Theorem 2.2, where again the $\mathcal{O}(\varepsilon)$-term is uniform in $y$. The strategy to establish sample paths estimates for (25) is to successively linearize both fast and slow variables around their deterministic counterpart (i.e. the solution for $\sigma_1 = \sigma_2 = 0$) denoted by $y_t^{\text{det}}, \bar{x}(y_t^{\text{det}}, \varepsilon)$. The deviations are then described by

$$
\begin{aligned}
\xi_t &= x_t - \bar{x}(y_t^{\text{det}} + \eta_t, \varepsilon), \\
\eta_t &= y_t - y_t^{\text{det}}.
\end{aligned}
\tag{26}
$$

They satisfy the following SDE, whose form is obtained by successively applying Taylor's theorem (int. always stands for the appropriate intermediate value)

$$
\begin{aligned}
\mathrm{d}\xi_t &= \frac{1}{\varepsilon} \left[ f(\xi_t + \bar{x}(y_t, \varepsilon), y_t, \varepsilon) - f(\bar{x}(y_t, \varepsilon), y_t, \varepsilon) \right] \mathrm{d}t + \frac{\sigma_1}{\varepsilon^{H_1}} \mathrm{d}W_t^{H_1} \\
&= \frac{1}{\varepsilon} \left[ \partial_x f(\bar{x}(y_t, \varepsilon), y_t, \varepsilon)\xi_t + \partial_{xx} f(\text{int.}, y_t, \varepsilon)\xi_t^2 \right] + \frac{\sigma_1}{\varepsilon^{H_1}} \mathrm{d}W_t^{H_1} \\
&= \frac{1}{\varepsilon} \left[ \tilde{a}(t, \varepsilon)\xi_t + \tilde{b}(\xi_t, \eta_t, t, \varepsilon) \right] + \frac{\sigma_1}{\varepsilon^{H_1}} \mathrm{d}W_t^{H_1},
\end{aligned}
$$

where

$$\tilde{a}(t, \varepsilon) = \partial_x f(\bar{x}(y_t^{\text{det}}, \varepsilon), y_t^{\text{det}}, \varepsilon) = a(t) + \mathcal{O}(\varepsilon),$$

$$
\begin{aligned}
b(\xi_t, \eta_t, t, \varepsilon) = \partial_{xx} f(\text{int.}, y_t, \varepsilon)\xi_t^2 &+ \left[ \partial_{xx} f(\bar{x}(\text{int.}, \varepsilon), \text{int.}, \varepsilon)\partial_y \bar{x}(\text{int.}, \varepsilon) \right. \\
&\left. + \partial_y f(\bar{x}(\text{int.}, \varepsilon), \text{int.}, \varepsilon) \right] \xi_t \eta_t,
\end{aligned}
$$

so that $|b(x, y, t, \varepsilon)| \le M_1(x^2 + |xy|)$, where $M_1$ is uniform in the variables due to the uniform boundedness assumption.

$$
\begin{aligned}
\mathrm{d}\eta_t &= \left[ g(x_t, y_t, \varepsilon) - g(\bar{x}(y_t, \varepsilon), y_t, \varepsilon) + g(\bar{x}(y_t, \varepsilon), y_t, \varepsilon) - g(\bar{x}(y_t^{\text{det}}, \varepsilon), y_t^{\text{det}}, \varepsilon) \right] \mathrm{d}t \\
&\quad + \sigma_2 \mathrm{d}W_t^{H_2} \\
&= \left[ \partial_x g(\bar{x}(y_t, \varepsilon), y_t, \varepsilon)\xi_t + \partial_{xx} g(\text{int.}, y_t, \varepsilon)\xi_t^2 \right. \\
&\quad + \left( \partial_x g(\bar{x}(y_t^{\text{det}}, \varepsilon), y_t^{\text{det}}, \varepsilon)\partial_y \bar{x}(y_t^{\text{det}}, \varepsilon) + \partial_y g(\bar{x}(y_t^{\text{det}}, \varepsilon), y_t^{\text{det}}, \varepsilon) \right) \eta_t \\
&\quad \left. + \tilde{R}(y_t, y_t^{\text{det}}, \varepsilon)\eta_t^2 \right] \mathrm{d}t + \sigma_2 \mathrm{d}W_t^{H_2} \\
&= \left[ c(y^{\text{det}}, \varepsilon)\xi_t + d(y^{\text{det}}, \varepsilon)\eta_t + R(\xi_t, \eta_t, \varepsilon) \right] \mathrm{d}t + \sigma_2 \mathrm{d}W_t^{H_2},
\end{aligned}
$$

where

$$c(y^{\mathrm{det}}, \varepsilon) = \partial_x g(\bar{x}(y_t^{\mathrm{det}}, \varepsilon), y_t^{\mathrm{det}}, \varepsilon),$$

$$d(y^{\mathrm{det}}, \varepsilon) = \partial_x g(\bar{x}(y_t^{\mathrm{det}}, \varepsilon), y_t^{\mathrm{det}}, \varepsilon)\partial_y \bar{x}(y_t^{\mathrm{det}}, \varepsilon) + \partial_y g(\bar{x}(y_t^{\mathrm{det}}, \varepsilon), y_t^{\mathrm{det}}, \varepsilon),$$

$$R(\xi_t, \eta_t, t, \varepsilon) = \partial_{xx} g(\mathrm{int.}, y_t, \varepsilon)\xi_t^2 + \tilde{R}(y_t, y_t^{\mathrm{det}}, \varepsilon)\eta_t^2 + \frac{\mathrm{d}}{\mathrm{d}y} c(\mathrm{int.}, \varepsilon)\xi_t \eta_t.$$

In particular the nonlinearity term satisfies for some $M_2 \geq 0$ (again uniform in the variables)

$$|R(x, y, t, \varepsilon)| \leq M_2(x^2 + y^2 + |xy|).$$

In order to prove that $x_t$ is concentrated in the neighborhood $\mathcal{B}(h) := \{(x, y) \in \mathbb{R}^2 : |x| < h\sqrt{\zeta(y)}\}$ around the slow manifold $C_\varepsilon$ with high probability define the exit times

$$\tau_{\mathcal{B}(h)} := \inf\{r > 0 : (\xi_r, y_r^{\mathrm{det}}) \notin \mathcal{B}(h)\},$$

$$\tau_{\eta, \tilde{h}} := \inf\{r > 0 : |\eta_r| > \tilde{h}\}.$$

Then we partition the event of $x_t$ in the following way

$$\mathbb{P}\left(\tau_{\mathcal{B}(h)} < t\right) = \mathbb{P}\left(\tau_{\mathcal{B}(h)} < t, \tau_{\eta, \tilde{h}} > \tau_{\mathcal{B}(h)}\right) + \mathbb{P}\left(\tau_{\mathcal{B}(h)} < t, \tau_{\eta, \tilde{h}} \leq \tau_{\mathcal{B}(h)}\right)$$

$$\leq \mathbb{P}\left(\tau_{\mathcal{B}(h)} < t \wedge \tau_{\eta, \tilde{h}}\right) + \mathbb{P}\left(\tau_{\eta, \tilde{h}} \leq t \wedge \tau_{\mathcal{B}(h)}\right).$$

Note that the first probability is of the form

$$\mathbb{P}\left(\tau_{\mathcal{B}(h)} < t \wedge \tau_{\eta, \tilde{h}}\right) = \mathbb{P}\left(\sup_{0 \leq r < t \wedge \tau_{\eta, \tilde{h}}} \frac{|\xi_r|}{\sqrt{\zeta(y_r^{\mathrm{det}})}} \geq h\right).$$

By the same technique used to prove Theorem 3.8 we get

$$\mathbb{P}\left(\sup_{0 \leq r < t} \frac{|\xi_r|}{\sqrt{\zeta(y_r^{\mathrm{det}})}} \geq h\right) \leq 2e \left\lceil \frac{|\alpha(t)|}{\varepsilon} \kappa^2 \frac{h^2}{\sigma^2}(1 + \mathcal{O}(\varepsilon)) \right\rceil \exp\left(-\kappa^2 \frac{h^2}{2\sigma^2}\right),$$

which is valid as long as $h\zeta^+ + \sqrt{\zeta^-}\tilde{h} \leq \frac{a\sqrt{\zeta^-}}{2M_1}$. It remains to estimate $\mathbb{P}\left(\tau_{\eta, \tilde{h}} \leq t \wedge \tau_{\mathcal{B}(h)}\right)$. This issue is however tightly linked to investigating the behavior of non-stable or even non-hyperbolic dynamics under fractional noise because we have no additional assumptions on the slow dynamics. In the event $\{\tau_{\eta, \tilde{h}} \leq t\}$ we conjecture that a reduction to the slow variables should be possible. The reduced equation is then given by

$$\mathrm{d}y_t = g(x(t, \varepsilon), y_t, \varepsilon)\mathrm{d}t + \sigma_2 \mathrm{d}W_t^{H_2},$$

which will be illustrated by the simulations presented at the end of Sect. 3.4.

## 3.4 Example

We consider the climate model analyzed in [4, Section 6.2.1]. It is a simple model describing the difference of temperature $\Delta T = T_1 - T_2$ and salinity $\Delta S = S_1 - S_2$ between low latitude $(T_1, S_1)$ and high latitude $(T_2, S_2)$ by a system of coupled differential equations

$$\frac{d}{ds}\Delta T = -\frac{1}{\tau_r}(\Delta T - \theta) - \left(\frac{1}{\tau_d} + \frac{q}{V}(\alpha_S \Delta S - \alpha_T \Delta T)^2\right)\Delta T,$$

$$\frac{d}{ds}\Delta S = \frac{F}{H}S_0 - \left(\frac{1}{\tau_d} + \frac{q}{V}(\alpha_S \Delta S - \alpha_T \Delta T)^2\right)\Delta S.$$

Here $\tau_r$ stands for the relaxation time of $\Delta T$ to its reference value $\theta$, $F$ is the freshwater flux, $H$ the depth of the ocean, $S_0$ a reference salinity. Furthermore, $\tau_d$ is the diffusion timescale, $q$ the Poiseuille transport coefficient and $V$ the volume of the box the system is contained in. The influence of external sources, internal fluctuations, and/or microscopic effects can be incorporated into the model via noise terms. For example, daily weather variations certainly influence the temperature $T$ and salinity $S$. Yet, a precise/detailed modelling of these terms would be far too expensive computationally and would make the model intractable analytically. We know, as discussed in the introduction to this work, that using white noise generally does not represent temperature fluctuations correctly but these are usually positively correlated. Since we have no further basic knowledge of the stochastic process it is quite natural to start by considering fBm with Hurst index $H > 1/2$. This allows us to model Gaussianity and positive correlations in time. After transforming our model into dimensionless variables $x = \Delta T/\theta$, $y = \alpha_S \Delta S/(\alpha_T \theta)$, rescaling time by $\tau_d$ and taking into consideration fractional noise with Hurst parameter $H > \frac{1}{2}$, this yields the system

$$
\begin{aligned}
dx_t &= \frac{1}{\varepsilon}\left[-(x_t - 1) - \varepsilon x_t(1 + \eta^2(x_t - y_t)^2)\right]dt + \frac{\sigma_1}{\varepsilon^H}dW_t^H,\\
dy_t &= \left[\mu - y_t(1 + \eta^2(x_t - y_t)^2)\right]dt + \sigma_2 dW_t^H,
\end{aligned}
\tag{27}
$$

where $\varepsilon = \tau_r/\tau_d$, $\eta^2 = \tau_d(\alpha_T \theta)^2 q/V$, and $\mu = F \cdot \alpha_S S_0 \tau_d/(\alpha_T \theta H)$. Note that the previous system is of the form (25). We consider the solution on a bounded time interval $[0, T]$, $T > 0$ to ensure the uniform boundedness of the corresponding functions, as imposed in Assumption 3.12.

The slow subsystem of the deterministic system is given by

$$0 = -(x_t - 1),$$

$$\frac{d}{dt}y_t = \dot{y}_t = \mu - y_t(1 + \eta^2(x_t - y_t)^2).$$

In particular, it has a normally hyperbolic critical manifold, namely

$$C_0 = \{(x, y) \in \mathbb{R}^2 : x = 1\},$$

which is even stable, as $\frac{d}{dx}(-(x-1)) = -1$. By Theorem 2.2 there exists an invariant manifold

$$C_\varepsilon = \{(x, y) \in \mathbb{R}^2 : x = \bar{x}(y, \varepsilon) := 1 + \mathcal{O}(\varepsilon)\}.$$

In order to apply the estimate from Theorem 3.5 note that $f(x, y, \varepsilon) = -(x_t - 1) - \varepsilon x_t(1 + \eta^2(x_t - y_t)^2)$, so that

$$\partial_x f(x, y, \varepsilon) = -1 - \varepsilon(1 + \eta^2(x - y)^2) - 2\varepsilon \eta^2(x - y).$$

Hence we have

$$a(y_t, \varepsilon) = \partial_x f(\bar{x}(y_t, \varepsilon), y_t, \varepsilon) = -1 + \mathcal{O}(\varepsilon).$$

By Proposition 3.2 there is an attracting slow manifold for the variance of the form

$$\zeta(t) = H\Gamma(2H) + \mathcal{O}(\varepsilon).$$

We conclude that, in the case that $y_t$ is deterministic, sample paths starting on the slow manifold $\bar{x}(y, \varepsilon)$ are concentrated in the set

$$\mathcal{B}(h) = \{(x, y) \in \mathbb{R}^2 : |x - 1 - \mathcal{O}(\varepsilon)| < h(H\Gamma(2H) + \mathcal{O}(\varepsilon))\},$$

or, more precisely, for $0 < t \leq T$ and initial data $(x_0, y_0) = (\bar{x}(y_0, \varepsilon), y_0)$

$$\mathbb{P}\left(\exists\, 0 \leq r \leq t : (x_r, y_r) \notin \mathcal{B}(h)\right) \leq 2e\left[\frac{t}{\varepsilon}\kappa\frac{h^2}{\sigma^2}(1 + \mathcal{O}(\varepsilon))\right]\exp\left(-\kappa\frac{h^2}{2\sigma^2}\right).$$

Figure 1 indicates that for small enough noise the dynamics around the slow manifold should be governed by the equation

$$\frac{\mathrm{d}}{\mathrm{d}t}y_t = \dot{y}_t = \mu - y_t(1 + \eta^2(1 - y_t)^2) + \sigma_2\mathrm{d}W_t^H. \tag{28}$$

In (28) $\eta^2$ is a fixed parameter, while $\mu$ is proportional to the freshwater flux. It can be hence treated as a slowly varying parameter compared to the rescaled salinity. By setting $X := y$ and $Y := \mu$ we obtain another fast slow system subject to some noise

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}X_t = \dot{X}_t &= Y_t - X_t(1 + \eta^2(1 - X_t)^2) + \sigma_2(t)\mathrm{d}W_t^H, \\
\frac{\mathrm{d}}{\mathrm{d}t}Y_t = \dot{Y}_t &= \varepsilon g(X_t, Y_t),
\end{aligned} \tag{29}$$

where $g \in \mathcal{C}^2(\mathbb{R}^2; \mathbb{R})$. In particular, we can apply our theory again on the two stable branches of the new slow manifold. We simulate now the reduced equation, similarly to [32, Section 7.1], where the same system was considered with respect to the Brownian motion. The results for two different Hurst parameters (i.e. $H = 0.6$ and $H = 0.8$) are illustrated in Figs. 2 and 3.

## 4 The Multi-dimensional Case

In this section, we make the first steps towards extending our theory to the multi-dimensional case. Note that we keep the same notation as for the one-dimensional objects. We start again with uniform slow dynamics and consider the fast-slow system system in slow time

$$\begin{aligned}
\mathrm{d}x_t &= \frac{1}{\varepsilon}f(x_t, y_t, \varepsilon)\mathrm{d}t + \frac{\sigma}{\varepsilon^H}\mathrm{d}W_t^H \\
\mathrm{d}y_t &= 1\mathrm{d}t,
\end{aligned} \tag{30}$$

under the following assumptions. Let $m \geq 1$.

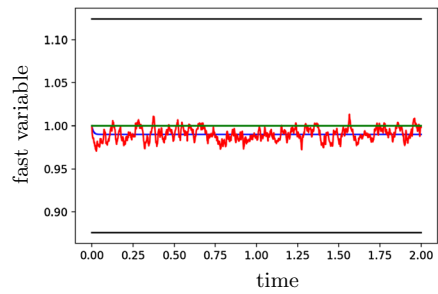**Assumption 4.1** Stable autonomous multi-dimensional case

1. *Regularity* The function $f \in \mathcal{C}^2(\mathbb{R}^m \times [0, \infty)^2; \mathbb{R})$, as well as its derivatives up to order 2 are uniformly bounded by a constant $M \geq 0$ on an interval $I = [0, \infty)$ or $I = [0, T]$, $T > 0$.
2. *Critical manifold* There is an $x^* : [0, \infty) \to \mathbb{R}^m$ such that
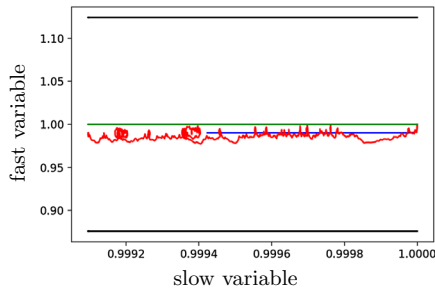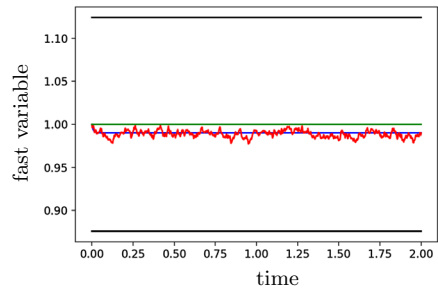
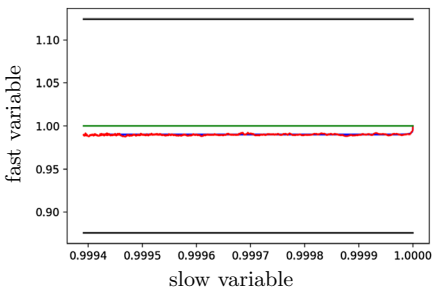$$f(x^*(t), t, 0) = 0$$

for all $t \in [0, \infty)$.
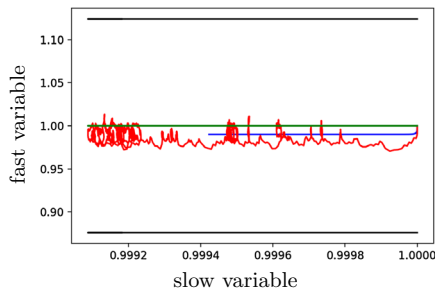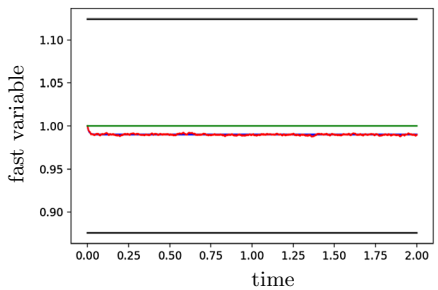
**Fig. 1** Equation (27) simulated for Hurst parameter $H = 0.7$, $\varepsilon = 0.01$, $\sigma_2 = 0$ and different $\sigma_1$. The stochastic solution is displayed red, the deterministic one is blue, the critical manifold is in green and the neighborhood $\mathcal{B}(h)$ for $h = 0.2$ is in black (Color figure online)

3. *Stability* The critical manifold is asymptotically stable, i.e. the Jacobian matrix

$$A(t) := \partial_x f(x^*(t), t, 0)$$

only contains eigenvalues with negative real part. In addition, its linearization is independent of time, i.e. $A(t) \equiv A$.

4. *Noise* $(W_t^H)_{t \geq 0}$ is an $m$-dimensional fractional Brownian motion.

These assumptions guarantee the existence and uniqueness of (30) due to Remark 2.12(R2). Furthermore, recall that under these assumptions there is a slow manifold

$\sigma_2 = 0.1,\ h = 3$

$\sigma_2 = 0.01,\ h = 1$

Periodic noise with $\sigma_2(t) = 0.05\sin(10t)+0.15$, $h = 3$

Process starting on $(1.2, 2)$ and moving backward in time. We also give a neighborhood for the second stable branch; $\sigma_2 = 0.1,\ h = 3$

**Fig. 2** Equation (29) simulated for Hurst parameter $H = 0.6$, $\varepsilon = 0.01$. and different noise. The stochastic solution is displayed red, the deterministic one is blue, the critical manifold is in green and the neighborhood $\mathcal{B}(h)$ for varying $h$ is in black (Color figure online)

$$C_\varepsilon = \left\{ (x, t) \in \mathbb{R}^m \times I : x = \bar{x}(t, \varepsilon) = x^*(t) + \mathcal{O}(\varepsilon) \right\},$$

due to Fenichel–Tikhonov (Theorem 2.2). We start again by examining the behavior of the linearized system around $C_\varepsilon$. For a solution $(x_t)_{t \in I}$ of (30) set $\xi_t := x_t - \bar{x}(t, \varepsilon)$, then $(\xi_t)_{t \in I}$ satisfies the equation

$$
\begin{aligned}
\mathrm{d}\xi_t &= \frac{1}{\varepsilon}\left[ f(\xi_t + \bar{x}(t, \varepsilon), t, \varepsilon) - f(\bar{x}(t, \varepsilon), t, \varepsilon) \right] \mathrm{d}t + \frac{\sigma}{\varepsilon^H} \mathrm{d}W_t^H \\
&= \frac{1}{\varepsilon}[A(t, \varepsilon)\xi_t + B(\xi_t, t, \varepsilon)]\mathrm{d}t + \frac{\sigma}{\varepsilon^H} \mathrm{d}W_t^H,
\end{aligned}
\tag{31}
$$

where

$$A(t, \varepsilon) = \partial_x f(\bar{x}(t, \varepsilon), t, \varepsilon) = \partial_x f(x^*(t), t, 0) + \mathcal{O}(\varepsilon),$$

$$\|B(x, t, \varepsilon)\|_2 \le M \|x\|_2,$$

with $\|\cdot\|_2$ being the operator norm with respect to the Euclidean norm. For simplicity, we analyze the linearization with $A$ being the drift term instead of $A + \mathcal{O}(\varepsilon)$, i.e. we consider

$$\mathrm{d}\xi_t = \frac{1}{\varepsilon}A\xi_t\mathrm{d}t + \frac{\sigma}{\varepsilon^H}\mathrm{d}W_t^H. \tag{32}$$

**Fig. 3** Equation (29) simulated for Hurst parameter $H = 0.8$, $\varepsilon = 0.01$. and different noise intensity. The stochastic solution is displayed red, the deterministic one is blue, the critical manifold is in green and the neighborhood $\mathcal{B}(h)$ for varying $h$ is in black (Color figure online)

The solution for $\xi_0 = 0$ $((x_t)_{t \in I}$ starting on the slow manifold $C_\varepsilon)$ is given by

$$\xi_t = \frac{\sigma}{\varepsilon^H} \int_0^t e^{A(t-u)} \mathrm{d} W_u^H. \tag{33}$$

Its covariance (matrix) can be computed as

$$\sigma^2 \, \Xi(t) := \mathrm{Cov}(\xi_t) = \frac{\sigma^2}{\varepsilon^{2H}} \int_0^t \int_0^t e^{A(t-u)} e^{A^\top (t-v)} H(2H-1) \, |u-v|^{2H-2} \, \mathrm{d} u \mathrm{d} v.$$

In the same way as in the one-dimensional case the rescaled covariance $t \mapsto \Xi(t)$ inherits the fast-slow structure.

**Proposition 4.2** *The so-called renormalized covariance $\Xi(t)$ satisfies the fast-slow ODE*

$$\varepsilon \frac{\mathrm{d}}{\mathrm{d} t} \Xi(t) = \varepsilon \dot{\Xi}(t) = A \Xi(t) + \Xi(t) A^\top + \frac{1}{\varepsilon^{2H-1}} \left[ P(t) + P(t)^\top \right], \tag{34}$$

*where*

$$P(t) = H(2H-1) \int_0^t e^{A(t-u)/\varepsilon} (t-u)^{2H-2} \mathrm{d} u.$$

*In particular, there is an (even globally) asymptotically stable slow manifold of the system of the form*

$$\bar{X}(t) = \int_0^\infty e^{Au} \left( Q(t) + Q(t)^\top \right) e^{A^\top u} \mathrm{d}u + \mathcal{O}(\varepsilon), \tag{35}$$

*where*

$$Q(t) = H(2H - 1) \int_0^\infty e^{Au} u^{2H-2} \mathrm{d}u.$$

**Proof** We again differentiate $t \mapsto \Xi(t)$ to obtain the ODE

$$\varepsilon \frac{\mathrm{d}}{\mathrm{d}t} \Xi(t) = \varepsilon \dot{\Xi}(t) = A\Xi(t) + \Xi(t)A^\top + \frac{1}{\varepsilon^{2H-1}} \left[ P(t) + P(t)^\top \right],$$

where

$$P(t) = H(2H - 1) \int_0^t e^{A(t-u)/\varepsilon} (t - u)^{2H-2} \mathrm{d}u.$$

In order to be able to take the singular limit $\varepsilon \to 0$ and apply Fenichel–Tikhonov (Theorem 2.2) we need to prove at least one times continuous differentiability in $\varepsilon = 0$. To do this, rewrite $\frac{1}{\varepsilon^{2H-1}} P(t)$ by substituting $v = \frac{t-u}{\varepsilon}$

$$H(2H - 1) \int_0^{\frac{t}{\varepsilon}} e^{Av} v^{2H-2} \mathrm{d}v \longrightarrow H(2H - 1) \int_0^\infty e^{Av} v^{2H-2} \mathrm{d}v \quad \text{for } \varepsilon \to 0. \tag{36}$$

This implies continuity in $\varepsilon = 0$. To see that the right hand side of (34) is continuously differentiable in $\varepsilon = 0$, it is sufficient to check it for the integral $P(t)$

$$\frac{\mathrm{d}}{\mathrm{d}\varepsilon} (P(t)) = -H(2H - 1) \int_0^t e^{A(t-u)/\varepsilon} A(t - u)/\varepsilon^2 (t - u)^{2H-2} \mathrm{d}u,$$

where the limit for $\varepsilon \to 0$ exists because the exponential term dominates the polynomial term in $\varepsilon$. The slow subsystem hence reads

$$0 = A\Xi(t) + \Xi(t)A^\top + \left[ Q(t) + Q(t)^\top \right],$$

where

$$Q(t) = H(2H - 1) \int_0^\infty e^{Au} u^{2H-2} \mathrm{d}u.$$

This is a Lyapunov equation, and according to Lemma 4.4 it has the unique solution

$$X^*(t) = \int_0^\infty e^{Au} \left( Q(t) + Q(t)^\top \right) e^{A^\top u} \mathrm{d}u.$$

By Fenichel–Tikhonov (Theorem 2.2) we conclude that there is an asymptotically stable manifold of the form

$$\bar{X}(t) = \int_0^\infty e^{Au} \left( Q(t) + Q(t)^\top \right) e^{A^\top u} \mathrm{d}u + \mathcal{O}(\varepsilon).$$

Note again that the stability property, which carries over from the critical manifold, is even global due to linearity of the ODE (34). □

**Remark 4.3** We need that linearization (32) is autonomous in this section for taking the singular limit in (36). In the non-autonomous case we need to compute the limit of $\Phi(t, t - \varepsilon v)$ for $\varepsilon \to 0$. We suspect that $\Phi(t, t - \varepsilon v) \to e^{A(t)v}$.

In order to investigate a multi-dimensional Lyapunov-Equation, we rely on the following result, see [4, Lemma 5.1.2].

**Lemma 4.4** (Lyapunov Equation) *Let $A \in \mathbb{R}^{p \times p}$ and $B \in \mathbb{R}^{q \times q}$ with eigenvalues $a_1, \ldots, a_p$ and $b_1, \ldots, b_q$. Then the operator $L : \mathbb{R}^{p \times q} \to \mathbb{R}^{p \times q}$ defined by*

$$LX = AX + XB$$

*has eigenvalues of the form $\{a_i + b_j\}_{i=1,\ldots,p,\, j=1,\ldots,q}$. In particular, $L$ is invertible if and only if $A$ and $-B$ don't have any common eigenvalue. Moreover, if all eigenvalues of $A$ and $B$ have negative real part, then for any $C \in \mathbb{R}^{p \times q}$ the unique solution of the so-called Lyapunov equation $AX + XB + C = 0$ is of the form*

$$X = \int_0^\infty e^{Au} C e^{Bu} \mathrm{d}u.$$

Note again that due to the linearity of the operator $LX = AX + XA^\top$ the rescaled covariance $t \mapsto \frac{1}{\sigma^2} \mathrm{Cov}(\xi_t)$ as solution of (34) with starting value $\frac{1}{\sigma^2} \mathrm{Cov}(\xi_0) = 0$ satisfies the following equation

$$\mathrm{Cov}(\xi_t) = \sigma^2 \left( \bar{X}(t) - e^{At/\varepsilon} \bar{X}(0) e^{A^\top t/\varepsilon} \right), \tag{37}$$

which explicitly depicts the exponentially fast approach of the covariance towards the slow manifold, as it could have been already concluded by Fenichel–Tikhonov (Theorem 2.2). This justifies the choice of our neighborhood this time, depending on the critical manifold

$$\mathcal{B}(h) = \left\{ (x, t) : t \in I, \langle x, X^*(t)^{-1} x \rangle < h^2 \right\}.$$

As already previously mentioned in Remark 3.4 choosing the neighborhood depending on the critical manifold instead of the slow manifold does not worsen our estimates. So we expect the same to be true in the higher dimensional case. Therefore, we have used the critical manifold $X^*$ (which is time-independent in our case) this time because our strategy depends on diagonalizing, and we do not spell out the additional technical details regarding the $\mathcal{O}(\varepsilon)$-term.

## 4.1 Estimates on the Deviations

### 4.1.1 No Restrictions on the Linearization

The proof of Theorem 3.7 can be immediately extended the multi-dimensional case by proving the mean-square Hölder continuity in each component of the covariance.

**Lemma 4.5** *Let $t \geq r_1 > r_2 \geq 0$, then there is a constant $G = G(t, \varepsilon, \sigma, H) > 0$ such that*

$$\left\| \mathbb{E} \left[ \left( \xi_{r_1} - \xi_{r_2} \right) \left( \xi_{r_1} - \xi_{r_2} \right)^\top \right] \right\|_2 \leq G \left| r_1 - r_2 \right|^{2H}.$$

**Proof** Let $t \geq r_1 > r_2 \geq 0$, then

$$\left\| \mathbb{E}\left[ (\xi_{r_1} - \xi_{r_2}) (\xi_{r_1} - \xi_{r_2})^\top \right] \right\|_2$$

$$= \frac{\sigma^2}{\varepsilon^{2H}} \left\| \int_0^{r_1} \int_0^{r_1} \left( e^{A(r_1-u)/\varepsilon} - \mathbb{1}_{\{u \leq r_2\}} e^{A(r_2-u)/\varepsilon} \right) \right.$$

$$\left. \cdot \left( e^{A^\top (r_1-v)/\varepsilon} - \mathbb{1}_{\{v \leq r_2\}} e^{A^\top (r_2-v)/\varepsilon} \right) \phi(u-v) \mathrm{d}u \mathrm{d}v \right\|_2$$

$$\leq \frac{\sigma^2}{\varepsilon^{2H}} \int_0^{r_2} \int_0^{r_2} \left\| e^{A(r_1-u)/\varepsilon} - e^{A(r_2-u)/\varepsilon} \right\|_2 \left\| e^{A^\top (r_1-v)/\varepsilon} - e^{A^\top (r_2-v)/\varepsilon} \right\|_2 \phi(u-v) \mathrm{d}u \mathrm{d}v \tag{38}$$

$$+ \frac{\sigma^2}{\varepsilon^{2H}} \int_{r_2}^{r_1} \int_0^{r_2} \left\| e^{A(r_1-u)/\varepsilon} - e^{A(r_2-u)/\varepsilon} \right\|_2 \left\| e^{A^\top (r_1-v)/\varepsilon} \right\|_2 \phi(u-v) \mathrm{d}u \mathrm{d}v \tag{39}$$

$$+ \frac{\sigma^2}{\varepsilon^{2H}} \int_0^{r_2} \int_{r_2}^{r_1} \left\| e^{A(r_1-u)/\varepsilon} \right\|_2 \left\| e^{A^\top (r_1-v)/\varepsilon} - e^{A^\top (r_2-v)/\varepsilon} \right\|_2 \phi(u-v) \mathrm{d}u \mathrm{d}v \tag{40}$$

$$+ \frac{\sigma^2}{\varepsilon^{2H}} \int_{r_2}^{r_1} \int_{r_2}^{r_1} \left\| e^{A(r_1-u)/\varepsilon} \right\|_2 \left\| e^{A^\top (r_1-v)/\varepsilon} \right\|_2 \phi(u-v) \mathrm{d}u \mathrm{d}v. \tag{41}$$

Since $A$ only has eigenvalues with negative real part we have for $r \geq u \geq 0$

$$\left\| e^{A(r-u)/\varepsilon} \right\|_2 \leq 1, \quad \left\| e^{A^\top (r-u)/\varepsilon} \right\|_2 \leq 1,$$

and for $r_1 \geq u \geq 0$, $r_2 \geq u \geq 0$

$$\left\| e^{A(r_1-u)/\varepsilon} - e^{A(r_2-u)/\varepsilon} \right\|_2 \leq \frac{a}{\varepsilon} |r_1 - r_2|, \quad \left\| e^{A^\top (r_1-u)/\varepsilon} - e^{A^\top (r_2-u)/\varepsilon} \right\|_2 \leq \frac{a}{\varepsilon} |r_1 - r_2|,$$

where $a := \max\{|\lambda| : \lambda \text{ eigenvalue of } A\}$. This enables us to prove the result similarly as in the one-dimensional case, i.e., a straightforward calculations using the last result now shows the required Hölder bounds by estimating (38)–(41). □

The mean-square Hölder continuity of $(\xi_t)_{t \in I}$ implies the same for each component. Hence, we can establish the following qualitative result.

**Theorem 4.6** *Let $t \in I$. Then under Assumption 4.1 there is a constant $K = K(t, \varepsilon, \sigma, H) > 0$ such that for $h > 0$ the following estimate holds true for $\varepsilon > 0$ small enough*

$$\mathbb{P}\left( \sup_{0 \leq r \leq t} \langle \xi_r, X^*(r)^{-1} \xi_r \rangle \geq h^2 \right) \leq \sum_{k=1}^m Kt \left( \sqrt{\lambda_k} h d_k^* \right)^{\frac{1}{H}} \exp\left( -\frac{\lambda_k h^2}{2\sigma^2} (1 - \mathcal{O}(\varepsilon)) \right),$$

*where $\lambda_k \geq 0$ with $\sum_{k=1}^m \lambda_k = 1$ and $d_k^*$ denote the (time-independent) eigenvalues of $X^*$.*

**Proof** Note that the critical manifold $X^*$ is symmetric and in the autonomous case it is time independent in addition. This implies that it is diagonalizable with respect to an orthogonal matrix $O$ (independent of time). Let $O^\top = (O_1^\top, \ldots, O_m^\top)$, where $O_k$ denotes the $k$-th row of $O$ and $d_k^*(r) \equiv d_k^*$ be the corresponding eigenvalues. This enables us to reduce the problem to the estimate of the one-dimensional problem, using the notation $D^*(r) = \mathrm{diag}(d_k^*)$,

$$\mathbb{P}\left(\sup_{0\le r\le t}\langle\xi_r,X^*(r)^{-1}\xi_r\rangle\ge h^2\right)=\mathbb{P}\left(\sup_{0\le r\le t}\langle O\xi_r,D^*(r)^{-1}O\xi_r\rangle\ge h^2\right)$$

$$=\mathbb{P}\left(\sup_{0\le r\le t}\sum_{k=1}^m O_k\xi_r d_k^*(r)^{-1}O_k\xi_r\ge h^2\right)$$

$$\le\sum_{k=1}^m\mathbb{P}\left(\sup_{0\le r\le t}d_k^*(r)^{-1}(O_k\xi_r)^2\ge\lambda_k h^2\right)$$

$$\le 2\sum_{k=1}^m\mathbb{P}\left(\sup_{0\le r\le t}\frac{O_k\xi_r}{\sqrt{d_k^*(r)}}\ge\lambda_k h^2\right)$$

$$=2\sum_{k=1}^m\mathbb{P}\left(\sup_{0\le r\le t}O_k\xi_r\ge\lambda_k h^2 d_k^*\right).$$

We have already proven that $(\xi_t)_{t\in I}$ is mean-square Hölder continuous in Lemma 4.5, which directly implies the same property for the components in the $O$-coordinate system. This means that we can apply Theorem 2.16 for $O_k\xi$. This leads to

$$\mathbb{P}\left(\sup_{0\le r\le t}O_k\xi_r\ge\lambda_k h^2 d_k^*\right)\le Kt\left(\sqrt{\lambda_k}h d_k^*\right)^{\frac{1}{H}}\exp\left(-\frac{\lambda_k h^2 d_k^*}{2\sup_{0\le r\le t}\mathrm{Var}(O_k\xi_r)}\right).$$

Now note that (37) written in the $k$-th component in the $O$-coordinate system reads as

$$\mathrm{Var}(O_k\xi_t)=\left(O\mathrm{Cov}(\xi_t)O^\top\right)_{kk}$$

$$=\sigma^2\left(O\left(X^*(t)+\mathcal{O}(\varepsilon)\right)O^\top-Oe^{At/\varepsilon}\bar X(0)e^{A^\top t/\varepsilon}O^\top\right)_{kk}$$

$$=\sigma^2\left((d_k^*(t)+\mathcal{O}(\varepsilon))-e^{2a_k t/\varepsilon}\left(O\bar X(0)O^\top\right)_{kk}\right)$$

$$=\sigma^2\left((d_k^*+\mathcal{O}(\varepsilon))-e^{2a_k t/\varepsilon}\left(O\bar X(0)O^\top\right)_{kk}\right),$$

This further implies

$$\mathbb{P}\left(\sup_{0\le r\le t}O_k\xi_r\ge\lambda_k h^2 d_k^*\right)\le Kt\left(\sqrt{\lambda_k}h d_k^*\right)^{\frac{1}{H}}\exp\left(-\frac{\lambda_k h^2 d_k^*}{2\sup_{0\le r\le t}\mathrm{Var}(O_k\xi_r)}\right)$$

$$\le Kt\left(\sqrt{\lambda_k}h d_k^*\right)^{\frac{1}{H}}\exp\left(-\frac{\lambda_k h^2}{2\sigma^2}\frac{d_k^*}{d_k^*+\mathcal{O}(\varepsilon)}\right)$$

$$\le Kt\left(\sqrt{\lambda_k}h d_k^*\right)^{\frac{1}{H}}\exp\left(-\frac{\lambda_k h^2}{2\sigma^2}(1-\mathcal{O}(\varepsilon))\right).$$

Summing over the dimensions yields

$$\mathbb{P}\left(\sup_{0\le r\le t}\langle\xi_r,X^*(r)^{-1}\xi_r\rangle\ge h^2\right)\le\sum_{k=1}^m Kt\left(\sqrt{\lambda_k}h d_k^*\right)^{\frac{1}{H}}\exp\left(-\frac{\lambda_k h^2}{2\sigma^2}(1-\mathcal{O}(\varepsilon))\right)$$

and this finishes the proof.                                                                                       □

In the case when $A$ is normal we get a nice description of the $d_k^*$. The corresponding $k$-th component of the diagonalized critical manifold results in

$$d_k^* = H(2H-1)\frac{1}{|a_k + \overline{a_k}|} \int_0^\infty \left( e^{a_k v} + e^{\overline{a_k} v} \right) v^{2H-2} dv,$$

which simplifies even further if all eigenvalues are real

$$d_k^* = H(2H-1)\frac{1}{|a_k|} \int_0^\infty e^{a_k v} v^{2H-2} dv$$

$$= H(2H-1)\frac{1}{|a_k|^{2H}} \Gamma(2H-1) = \frac{1}{|a_k|^{2H}} H\Gamma(2H).$$

### 4.1.2 Symmetric Linearization

From now on, we consider the case when $A$ is a symmetric matrix (i.e. $A = A^\top$). The reason for this restriction is that in the following the proves to bound the probability of $(\xi_t)_{t \in I}$ exiting the neighborhood up to time $t$

$$\mathbb{P}\left( \exists\, 0 \le r \le t :\ (x_r, y_r) \notin \mathcal{B}(h) \right)$$

is based on linearizing the underlying system and understanding the structure of the eigenvalues of the covariance. We actually require normality of $A$ for this strategy to work as it is sufficient to use the functional equality of the matrix exponential. Furthermore, $e^A$ inherits the normality structure, which is a necessary and sufficient criterion to characterize the eigenvalues of $e^{A\lambda} e^{A^\top \mu}$, $\lambda, \mu \ge 0$. To be able to generalize the result of variant 1 it is crucial that the eigenvalues of $A$ are all real. These two criteria already imply that $A$ is symmetric. In particular, we see that the critical manifold is of the form

$$X^*(t) = H(2H-1) \int_0^\infty e^{Au} \left( \int_0^\infty \left( e^{Av} + e^{A^\top v} \right) v^{2H-2} dv \right) e^{A^\top u} du$$

$$= H(2H-1) \int_0^\infty \int_0^\infty \left( e^{A(u+v)+A^\top u} + e^{Au+A^\top(v+u)} \right) v^{2H-2} dv du.$$

Thanks to the discussion above the we can consider the diagonalization $D^*(t) := UX^*(t)U^\top$. Its $k$-th ($k = 1, \ldots, m$) diagonal component is given by

$$d_k^*(t) := \left( UX^*(t)U^\top \right)_{kk} = H(2H-1) \int_0^\infty \int_0^\infty \left( e^{a_k(u+v)+a_k u} + e^{a_k u+a_k(v+u)} \right) v^{2H-2} dv du$$

$$= H(2H-1)\frac{1}{|a_k|} \int_0^\infty e^{a_k v} v^{2H-2} dv$$

$$= H(2H-1)\frac{1}{|a_k|^{2H}} \Gamma(2H-1) = \frac{1}{|a_k|^{2H}} H\Gamma(2H).$$

Similarly we can rewrite the covariance

$$\mathrm{Cov}(\xi_t) = \frac{\sigma^2}{\varepsilon^{2H}} \int_0^t \int_0^t e^{A(t-u)} e^{A^\top(t-v)} H(2H-1) |u-v|^{2H-2} \, du dv$$

$$= \frac{\sigma^2}{\varepsilon^{2H}} H(2H-1) \int_0^t \int_0^t e^{A(t-u)+A^\top(t-v)} |u-v|^{2H-2} \, du dv.$$

and diagonalize it with respect to the same $U$ and consider its $k$-th ($k = 1, \ldots, m$) component

$$\left(U\mathrm{Cov}(\xi_t)U^\top\right)_{kk} = (\mathrm{Cov}(U\xi_t))_{kk}$$
$$= \frac{\sigma^2}{\varepsilon^{2H}} H(2H-1) \int_0^t \int_0^t e^{a_k(t-u)+a_k(t-v)} |u-v|^{2H-2} \,\mathrm{d}u\mathrm{d}v.$$

We obtain the following result

**Theorem 4.7** *Let $t \in I$. Then under Assumption* 4.1 *and if $A$ is in addition symmetric with real eigenvalues $a_1, \ldots, a_m$ there is a constant such that for $h > 0$ the following estimate holds for $\varepsilon > 0$ small enough*

$$\mathbb{P}\left(\sup_{0\leq r\leq t} \langle \xi_r, X^*(r)^{-1}\xi_r\rangle \geq h^2\right) 2e\left\lceil \frac{|a_+|t}{\varepsilon} \frac{h^2}{\sigma^2}(1+\mathcal{O}(\varepsilon))\right\rceil \exp\left(-\frac{h^2}{2\sigma^2 m}\right),$$

*where $a_+ = \max\{|\mu| : \mu$ eigenvalue of $A\}$.*

**Proof** Let $U^\top = \left(U_1^\top \cdots U_m^\top\right)$, where $U_k$ denotes the $k$-th row of $U$. Now

$$\mathbb{P}\left(\sup_{0\leq r\leq t} \langle \xi_r, X^*(r)^{-1}\xi_r\rangle \geq h^2\right) = \mathbb{P}\left(\sup_{0\leq r\leq t} \langle U\xi_r, D^*(r)^{-1}U\xi_r\rangle \geq h^2\right)$$
$$= \mathbb{P}\left(\sup_{0\leq r\leq t} \sum_{k=1}^m (U_k\xi_r)d_k^*(r)^{-1}(U_k\xi_r) \geq h^2\right)$$

for $\lambda_k \geq 0$ with $\sum_{k=1}^m \lambda_k = 1$

$$\leq \sum_{k=1}^m \mathbb{P}\left(\sup_{0\leq r\leq t} d_k^*(r)^{-1}(U_k\xi_r)^2 \geq \lambda_k h^2\right)$$
$$= 2\sum_{k=1}^m \mathbb{P}\left(\sup_{0\leq r\leq t} \frac{U_k\xi_r}{\sqrt{d_k^*(r)}} \geq \sqrt{\lambda_k}h\right).$$

Due to the normality of $e^{A(t-u)}$ (inherited by $A$) the Gaussian process $U_k\xi$ has variance

$$\mathrm{Var}(U_k\xi_r) = \frac{\sigma^2}{\varepsilon^{2H}} \int_0^t \int_0^t e^{a_k(t-u)/\varepsilon} e^{a_k(t-v)/\varepsilon} H(2H-1) |u-v|^{2H-2} \,\mathrm{d}u\mathrm{d}v.$$

In particular, we can show that the process $\left(e^{-a_k t}U_k\xi_t\right)_{t\geq 0}$ satisfies the assumptions of Lemma 2.15, so that we can apply the Bernstein-type inequality. (The proof is completely analogous as in the one-dimensional case, see proof of Theorem 3.5.) To get a relation between the $k$-th value of the diagonalized covariance and the corresponding component of the critical manifold consider (37) in the $U$-coordinate system

$$\mathrm{Var}(U_k\xi_t) = \left(U\mathrm{Cov}(\xi_t)U^*\right)_{kk}$$
$$= \sigma^2\left(U\left(X^*(t) + \mathcal{O}(\varepsilon)\right)U^* - Ue^{At/\varepsilon}\bar{X}(0)e^{A^\top t/\varepsilon}U^*\right)_{kk}$$
$$= \sigma^2\left((d_k^*(t) + \mathcal{O}(\varepsilon)) - e^{2a_k t/\varepsilon}\left(U\bar{X}(0)U^*\right)_{kk}\right)$$
$$= \sigma^2\left((d_k^* + \mathcal{O}(\varepsilon)) - e^{2a_k t/\varepsilon}\left(U\bar{X}(0)U^*\right)_{kk}\right), \tag{42}$$

as $d^*(t) \equiv d^*$ is actually independent of time in our case. Now, we can use the same strategy as variant 1 in the one dimensional case for each $k$. For $\gamma \in (0, 1/2)$ let $0 = t_0 < t_1 < \ldots < t_N$ be a partition containing the interval $[0, t]$ such that

$$-a_k(t_{i+1} - t_i) = \varepsilon\gamma \quad \text{for} \quad 0 \le i < N = \left\lceil \frac{|a_k t|}{\varepsilon\gamma} \right\rceil.$$

We start by estimating the probability of the exit time on $[t_i, t_{i+1})$ for $i \in \{0, \ldots, N-1\}$

$$\mathbb{P}\left( \sup_{t_i \le r < t_{i+1}} \frac{U_k \xi_r}{\sqrt{d_k^*(r)}} \ge \sqrt{\lambda_k} h \right)$$

$$\le \mathbb{P}\left( \sup_{t_i \le r < t_{i+1}} e^{-a_k r/\varepsilon} U_k \xi_r \ge \sqrt{\lambda_k} h \inf_{t_i \le r < t_{i+1}} e^{-a_k r/\varepsilon} \sqrt{d_k^*(r)} \right).$$

Applying Lemma 2.15

$$\le \exp\left( -\frac{1}{2} \frac{\lambda_k h^2 \inf_{t_i \le r < t_{i+1}} e^{-2a_k r/\varepsilon} d_k^*(r)}{e^{-2a_k t_{i+1}/\varepsilon} \mathrm{Cov}(U_k \xi_{t_{i+1}})} \right)$$

$$\overset{(42)}{\le} \exp\left( -\frac{\lambda_k h^2}{2\sigma^2} \frac{d_k^*}{(d_k^* + \mathcal{O}(\varepsilon))} e^{2a_k(t_{i+1} - t_i)/\varepsilon} \right)$$

$$\le \exp\left( -\frac{\lambda_k h^2}{2\sigma^2} e^{-2\gamma} (1 - \mathcal{O}(\varepsilon)) \right).$$

Taking the union of the events that $U_k \xi$ has exited $\mathcal{B}(h)$ in $[t_i, t_{i+1})$ and using the subadditivity of the probability measure, yields

$$\mathbb{P}\left( \sup_{0 \le r \le t} \frac{U_k \xi_r}{\sqrt{d_k^*(r)}} \ge \sqrt{\lambda_k} h \right) \le \sum_{i=0}^{N-1} \mathbb{P}\left( \sup_{t_i \le r < t_{i+1}} \frac{U_k \xi_r}{\sqrt{d_k^*(r)}} \ge \sqrt{\lambda_k} h \right)$$

$$\le \left\lceil \frac{|a_k| t}{\varepsilon\gamma} \right\rceil \exp\left( -\frac{\lambda_k h^2}{2\sigma^2} e^{-2\gamma} (1 - \mathcal{O}(\varepsilon)) \right)$$

$$\le \left\lceil \frac{|a_k| t}{\varepsilon\gamma} \right\rceil \exp\left( \frac{\lambda_k h^2}{\sigma^2} \gamma(1 + \mathcal{O}(\varepsilon)) \right) \exp\left( -\frac{\lambda_k h^2}{2\sigma^2} \right).$$

Finding the minimal bound with respect to $\gamma$ now corresponds to optimizing

$$\gamma \mapsto \frac{|a_k| t}{\varepsilon\gamma} \exp\left( \frac{\lambda_k h^2}{\sigma^2} \gamma(1 + \mathcal{O}(\varepsilon)) \right) \exp\left( -\frac{\lambda_k h^2}{2\sigma^2} \right),$$

due to the monotonicity of $\lceil \cdot \rceil$. The optimal value is achieved for

$$\gamma = \frac{\sigma^2}{\lambda_k h^2} (1 + \mathcal{O}(\varepsilon))^{-1}.$$

Plugging this in the estimate gives the bound for the $k$-th component

$$\mathbb{P}\left( \sup_{0 \le r \le t} \frac{U_k \xi_r}{\sqrt{d_k^*(r)}} \ge \sqrt{\lambda_k} h \right) \le e \left\lceil \frac{|a_k| t}{\varepsilon} \frac{\lambda_k h^2}{\sigma^2} (1 + \mathcal{O}(\varepsilon)) \right\rceil \exp\left( -\frac{\lambda_k h^2}{2\sigma^2} \right).$$

Summing over the dimensions

$$
\mathbb{P}\left(\sup_{0 \le r \le t} \langle \xi_r, X^*(r)^{-1}\xi_r \rangle \ge h^2\right)
$$

$$
\le 2 \sum_{k=1}^{m} \mathbb{P}\left(\sup_{0 \le r \le t} \frac{U_k \xi_r}{\sqrt{d_k^*(r)}} \ge \sqrt{\lambda_k} h\right)
$$

$$
\le 2e \sum_{k=1}^{m} \left\lceil \frac{|a_k| \, t}{\varepsilon} \frac{\lambda_k h^2}{\sigma^2}(1 + \mathcal{O}(\varepsilon)) \right\rceil \exp\left(-\frac{\lambda_k h^2}{2\sigma^2}\right)
$$

$$
\le 2e \sum_{k=1}^{m} \left\lceil \frac{|a_+| \, t}{\varepsilon} \frac{\lambda_k h^2}{\sigma^2}(1 + \mathcal{O}(\varepsilon)) \right\rceil \exp\left(-\frac{\lambda_k h^2}{2\sigma^2}\right),
$$

where $a_+ := \max\{|\mu| : \mu \text{ eigenvalue of } A\}$. The optimal value is now attained by choosing $\lambda_k = \frac{1}{m}$. This yields

$$
\mathbb{P}\left(\sup_{0 \le r \le t} \langle \xi_r, X^*(r)^{-1}\xi_r \rangle \ge h^2\right) \le 2e \left\lceil \frac{|a_+| \, t}{\varepsilon} \frac{h^2}{\sigma^2}(1 + \mathcal{O}(\varepsilon)) \right\rceil \exp\left(-\frac{h^2}{2\sigma^2 m}\right).
$$

The proof is complete. $\qquad\square$

Due to the symmetry of $A$ we could have diagonalized the SDE in the beginning (i.e. look at it in the $U$-coordinate system) and done the whole theory established in Chapter 2 to get the existence of a slow manifold $\zeta_k(t)$ for $\mathrm{Var}(U_k \xi_t)$, which is of the form

$$
\zeta_k(t) = \frac{1}{|a_k|^{2H}} H\Gamma(2H) + \mathcal{O}(\varepsilon)
$$

$$
= d_k^* + \mathcal{O}(\varepsilon).
$$

However, we decided to use the results on the higher dimensional systems as much as possible to clearly indicate which steps of the proof can be generalized to more general classes of matrices beyond symmetric ones.

## 5 Outlook

This work provides a first step towards the investigation of fast-slow systems driven by fBm using sample paths estimates. So far we have examined the behavior close to a normally hyperbolic attracting invariant manifold in finite dimensions. Numerous extensions could be considered as next steps.

Having covered the uniformly attracting case, it is then natural to conjecture that there are scaling laws for the fluctuations as fast subsystem bifurcation points are approached, i.e., when hyperbolicity is lost. These results are available in the fast-slow Brownian motion case [32]. However, even when the fast dynamics is dominated by nonlinear terms [46] or one considers fast-slow maps with bounded noise [34] using modified proofs and additional technical tools it is possible to save many results. This robustness of the scaling laws near the loss of normal hyperbolicity leads one to conjecture that it will still be possible to prove such results for the fast-slow fBm case when $H \in (1/2, 1)$.

However, the analysis of fast-slow systems for $H \in (0, 1/2)$ is expected to be more complicated due to several reasons. First of all, a different integration theory has to be

considered, see for instance [5,11]. Furthermore, the kernel (8) we have used to develop an approximation of the variance by means of the slow manifold has a non-integrable singularity for $H < 1/2$. Last but not least, Bernstein-type inequalities as established in Lemma 2.15 do not hold true anymore, since the covariance function of the fractional Brownian motion is negative. Consequently, one has to develop completely different techniques in this case. Another related extension would be to analyze the dynamics of fast-slow systems driven by multiplicative noise. This issue, however, requires a more general theory than Itô-calculus because the fractional Brownian motion is not a semi-martingale.

Furthermore, one could consider other stochastic processes with memory. More precisely, one could think of other stochastic processes whose covariance functions are represented by

$$\int\limits_{0}^{\min\{s,t\}} K(s,r)K(t,r)\,\mathrm{d}r, \quad \text{for } s,t \geq 0,$$

for suitable square integrable kernels $K$, recall (5). Beyond fBm, further examples in this sense are the multi-fractional Brownian motion or the Rosenblatt process [9]. However, the analysis of fast-slow systems in this case is a challenging question, since these processes do not have in general stationary increments and are no longer Gaussian (as e.g. Rosenblatt processes).

Finally, one can also broaden the scope of the applications. Although climate dynamics is certainly a very important topic, where time-correlated noise is well-motivated by data such as temperature measurements, it is not the only possible application. Other areas, where fast-slow systems with fBm could be considered are financial markets. For example, assets could be modelled as fast variables influenced by fBm stochastic forcing, while the slow variables are political/social factors influencing the market, which change on a much slower time scale in many cases. Similar remarks and examples of concrete applications are likely also exist in many contexts in neuroscience, ecology and epidemiology, where stochastic fast-slow systems with Brownian motion are already used frequently.

## A On the Limit Superior of Gaussian Processes

The following proof has been developed in personal communication with Professor Andrey Dorogovtsev.

**Theorem A.1** *Let $(Y_t)_{t\geq 0}$ be a centered Gaussian process with covariance function $R(t,s) = \mathbb{E}[Y_t Y_s]$ satisfying*

1. $R(t,s) \to 0$ *for $t - s \to \infty$,*
2. $R(t,t) = 1$ *for all $t \geq 0$.*

*Then*

$$\limsup_{t\to\infty} Y_t = \infty \quad a.s.$$

**Proof** We aim to construct a sequence of independent random variables inductively to apply the Borel–Cantelli lemma. The strategy is highly based on the fact that for Gaussian random variables independence is equivalent to zero covariance. Let $t_1 := 0$. Given $t_1 < \cdots < t_n$ we apply Gram-Schmidt orthogonalization to the variables $Y_{t_1}, \ldots, Y_{t_n}$ in $L^2(\Omega, \mathcal{F}, \mathbb{P})$ to obtain uncorrelated and normalized random variables $S_1, \ldots, S_n$ satisfying for each $1 \leq i \leq n$

$$S_i = \sum_{j=1}^n a_{ij} Y_{t_j}, \quad Y_{t_i} = \sum_{j=1}^n b_{ij} S_j,$$

where the coefficients $a_{ij}$ and $b_{ij}$ only depend on the on the values $R(t_i, t_j)$ for $1 \leq i, j \leq n$ by construction. Now for any $t > t_n$ we can project $Y_t$ on the space $\mathrm{span}_{1\leq j\leq n}\{Y_{t_j}\} = \mathrm{span}_{1\leq j\leq n}\{S_j\}$

$$Y_t = \sum_{j=1}^n \mathbb{E}[Y_t S_j] S_j + Y_t'.$$

Observe that by construction $Y_t'$ is uncorrelated, and hence independent, of the the variables $Y_{t_1}, \ldots, Y_{t_n}$. This holds in particular if we choose $t_{n+1} \geq \max\{t_n, n+1\}$ with

$$\mathbb{E}\left[\left(\sum_{j=1}^{n+1} \mathbb{E}[Y_{t_{n+1}} S_j] S_j\right)^2\right] \leq \frac{1}{2^n},$$

where the latter is possible because $R(t,s) \to 0$ for $t - s \to \infty$. Inductively we get for each $n \in \mathbb{N}$

$$Y_{t_n} = Y_{t_n}' + \gamma_n.$$

For the sequence of independent random variables $(Y_{t_n}')_{n\in\mathbb{N}}$

$$\sum_{j=1}^\infty \mathbb{P}\left(Y_{t_n}' > c\right) \geq \sum_{j=1}^\infty \mathbb{P}\left(Y_{t_n} > c\right) + \mathbb{P}\left(-\gamma_n > c\right)$$

$$= \sum_{j=1}^\infty \mathbb{P}\left(Y_{t_1} > c\right) + \mathbb{P}\left(-\gamma_n > c\right)$$

$$= \infty,$$

because $\mathrm{Var}(Y_{t_n}) = R(t_n, t_n) = 1$ by assumption. By the Borel-Cantelli lemma we obtain

$$\limsup_{n\to\infty} Y_{t_n}' = \infty \quad a.s.$$

Now note that $\gamma_n \to 0$ as $n \to \infty$ as for any $c > 0$

$$\mathbb{P}\left(|\gamma_n| > c\right) \leq \frac{\mathbb{E}\left[|Y_{t_n} - Y'_{t_n}|^2\right]}{c^2}$$

$$= \frac{\mathbb{E}\left[\left(\sum_{j=1}^{n} \mathbb{E}[Y_{t_n} S_j] S_j\right)^2\right]}{c^2}$$

$$\leq \frac{1}{2^{n-1} c^2} \longrightarrow 0, \text{ as } n \to \infty.$$

This implies that

$$\limsup_{n \to \infty} Y_{t_n} = \limsup_{n \to \infty} (Y'_{t_n} + \gamma_n) = \infty \text{ a.s}$$

as required. □

Now we can apply this result to our setting.

**Corollary A.2** *Under the Assumptions of Theorem* 3.5 *we have*

$$\limsup_{t \to \infty} e^{-\alpha(t)/\varepsilon} \xi_t = \infty \text{ a.s.}$$

**Proof** Define $Z_t := \frac{e^{-\alpha(t)/\varepsilon} \xi_t}{\sqrt{\text{Var}(e^{-\alpha(t)/\varepsilon} \xi_t)}}$. Then $\text{Var}(Z_t) = \mathbb{E}(Z_t Z_t) = 1$ by construction. To prove the second assumption of Theorem A.1, observe that

$$R(t, s) = \mathbb{E}\left[\frac{\xi_t}{\sqrt{\text{Var}(\xi_t)}} \frac{\xi_s}{\sqrt{\text{Var}(\xi_s)}}\right].$$

Now note that $\text{Var}(\xi_t)$ is bounded for $t \in [0, \infty)$. This means it suffices to prove for every $s > 0$

$$\mathbb{E}[\xi_t \xi_s] \to 0 \text{ as } t \to 0.$$

The correlation function is given by

$$\mathbb{E}[\xi_t \xi_s]$$

$$= \frac{\sigma^2}{\varepsilon^{2H}} \int_0^s \int_0^t e^{\alpha(t,u)/\varepsilon} e^{\alpha(s,v)/\varepsilon} F(u) F(v) H(2H - 1) |u - v|^{2H-2} \, \mathrm{d}u \mathrm{d}v$$

$$\leq \frac{\sigma^2 F_+^2}{\varepsilon^{2H}} \int_0^s \int_0^t e^{\alpha(t,u)/\varepsilon} e^{\alpha(s,v)/\varepsilon} H(2H - 1) |u - v|^{2H-2} \, \mathrm{d}u \mathrm{d}v,$$

where $F_+ := \sup_{r \in [0,\infty)} F(r)$. Let $\varepsilon > 0$. Then there is $T > 0$ such that for $a := \sup_{r \in [0,\infty)} a(r)$

$$H(2H - 1) \frac{\sigma^2 F_+^2}{\varepsilon^{2H}} |T|^{2H-2} \frac{1}{|a|^2} < \frac{\varepsilon}{2}.$$

Now choose $t_0 > s + T$ such that for all $t \geq t_0$

$$\frac{\sigma^2 F_+^2}{\varepsilon^{2H}} \int_0^s \int_0^t e^{\alpha(t,u)/\varepsilon} e^{\alpha(s,v)/\varepsilon} H(2H - 1) |u - v|^{2H-2} \, \mathrm{d}u \mathrm{d}v$$

$$= \frac{\sigma^2 F_+^2}{\varepsilon^{2H}} e^{\alpha(t,s+T)} \underbrace{\int_0^s \int_0^{s+T} e^{\alpha(s+T,u)/\varepsilon} e^{\alpha(s,v)/\varepsilon} H(2H - 1) |u - v|^{2H-2} \, \mathrm{d}u \mathrm{d}v}_{\text{independent of } t} < \frac{\varepsilon}{2},$$

where we have used the semi-group property of $e^{\alpha(t,u)/\varepsilon}$. Putting this together we get for all $t \geq t_0$

$$\mathbb{E}[\xi_t \xi_s]$$

$$\leq \frac{\sigma^2 F_+^2}{\varepsilon^{2H}} \int_0^s \int_0^t e^{\alpha(t,u)/\varepsilon} e^{\alpha(s,v)/\varepsilon} H(2H-1) |u-v|^{2H-2} \, \mathrm{d}u \mathrm{d}v$$

$$= \frac{\sigma^2 F_+^2}{\varepsilon^{2H}} \int_0^s \int_0^{s+T} e^{\alpha(t,u)/\varepsilon} e^{\alpha(s,v)/\varepsilon} H(2H-1) |u-v|^{2H-2} \, \mathrm{d}u \mathrm{d}v$$

$$+ \frac{\sigma^2 F_+^2}{\varepsilon^{2H}} \int_0^s \int_{s+T}^t e^{\alpha(t,u)/\varepsilon} e^{\alpha(s,v)/\varepsilon} H(2H-1) |u-v|^{2H-2} \, \mathrm{d}u \mathrm{d}v$$

$$< \frac{\varepsilon}{2} + H(2H-1) \frac{\sigma^2 F_+^2}{\varepsilon^{2H}} |T|^{2H-2} \int_{s+T}^t e^{\alpha(t,u)/\varepsilon} \mathrm{d}u \int_0^s e^{\alpha(s,v)/\varepsilon} \mathrm{d}v$$

$$\leq \frac{\varepsilon}{2} + H(2H-1) \frac{\sigma^2 F_+^2}{\varepsilon^{2H}} |T|^{2H-2} \frac{1}{|a|^2} < \varepsilon.$$

Now, Theorem A.1 proves the claim. □

**Remark A.3** Note that one cannot directly apply classical probabilistic results such as the Borel-Cantelli Lemma, law of iterated logarithm or ergodic theorems to prove Corollary A.2, since the process $\xi$ has neither stationary nor independent increments.

# References

1. Ashkenazy, Y., Baker, D.R., Gildor, H., Havlin, S.: Nonlinearity and multifractality of climate change in the past 420,000 years. Geophys. Res. Lett. (2003). https://doi.org/10.1029/2003GL018099
2. Barboza, L., Li, B., Tingley, M.P., Viens, F.G.: Reconstructing past temperatures from natural proxies and estimated climate forcings using short-and long-memory models. Ann. Appl. Stat. **8**(4), 1966–2001 (2014)
3. Berglund, N., Gentz, B.: The effect of additive noise on dynamical hysteresis. Nonlinearity **15**(3), 605–632 (2002)
4. Berglund, N., Gentz, B.: Noise-Induced Phenomena in Slow-Fast Dynamical Systems. Springer, London (2006)
5. Biagini, F., Hu, Y., Øksendal, B., Zhang, T.: Stochastic Calculus for Fractional Brownian Motion and Applications. Springer, London (2008)
6. Blender, R., Fraedrich, K.: Long time memory in global warming simulations. Geophys. Res. Lett. (2003). https://doi.org/10.1029/2003GL017666
7. Bogachev, V.I.: Measure theory. Number Bd. 1 in measure theory. Springer, Berlin (2007)
8. Chekroun, M.D., Liu, H., Wang, S.: Stochastic parameterizing manifolds and non-Markovian reduced equations: stochastic manifolds for nonlinear SPDEs II. Springer, New York (2014)
9. Čoupek, P., Maslowski, B.: Stochastic evolution equations with Volterra noise. Stochastic Process. Appl. **127**(3), 877–900 (2017)
10. Davidsen, J., Griffin, J.: Volatility of unevenly sampled fractional Brownian motion: an application to ice core records. Phys. Rev. E **81**(1), 016107 (2010)
11. Decreusefond, L., et al.: Stochastic analysis of the fractional Brownian motion. Potential Anal. **10**(2), 177–214 (1999)
12. Decreusefond, L., Nualart, D.: Hitting times for Gaussian processes. Ann. Probab. **36**(1), 319–330 (2008)
13. Dijkstra, H.A.: Nonlinear Climate Dynamics. CUP, (2013)
14. Eichner, J.F., Koscielny-Bunde, E., Bunde, A., Havlin, S., Schellnhuber, H.J.: Power-law persistence and trends in the atmosphere: a detailed study of long temperature records. Phys. Rev. E **68**(4), 046133 (2003)
15. Fenichel, N.: Geometric singular perturbation theory for ordinary differential equations. J. Differ. Equat. **31**(1), 53–98 (1979)
16. Franzke, C.L., O'Kane, T.J., Berner, J., Williams, P.D., Lucarini, V.: Stochastic climate theory and modeling. Wiley Interdiscip. Rev. **6**(1), 63–78 (2015)

17. Gao, F.-Y., Kang, Y.-M., Chen, X., Chen, G.: Fractional Gaussian noise-enhanced information capacity of a nonlinear neuron model with binary signal input. Phys. Rev. E **97**(5), 052142 (2018)
18. Gardiner, C.: Stochastic methods, 4th edn. Springer, Berlin (2009)
19. Garrido-Atienza, M.J., Lu, K., Schmalfuss, B.: Local pathwise solutions to stochastic evolution equations driven by fractional Brownian motions with Hurst parameter $H \in (1/3, 1/2]$. Discret. Cont. Dyn. B **20**, 11 (2014)
20. Gnann, M., Kuehn, C., Pein, A.: Towards sample path estimates for fast-slow SPDEs. Euro. J. Appl. Math. **30**(5), 1004–1024 (2019)
21. Gubinelli, M., Lejay, A., Tindel, S.: Young integrals and SPDEs. Potential Anal. **25**(4), 307–326 (2006)
22. Gubinelli, M., Tindel, S., et al.: Rough evolution equations. Ann. Probab. **38**(1), 1–75 (2010)
23. Guckenheimer, J., Holmes, P.: Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields. Springer, New York (1983)
24. Hairer, M., Li, X.-M.: Averaging dynamics driven by fractional Brownian motion. Ann. Probab. (2019) (to appear)
25. Hek, G.: Geometric singular perturbation theory in biological practice. J. Math. Biol. **60**(3), 347–386 (2010)
26. Hesse, R., Neamṭu, A.: Local mild solutions for rough stochastic partial differential equations. J. Differential Equat. **267**(11), 6480–6538 (2019)
27. Jones, C.K.R.T.: Geometric Singular Perturbation Theory. In: Dynamical Systems (Montecatini Terme, 1994), volume 1609 of Lecture Notes in Mathematics, pp 44–118. Springer, (1995)
28. Kaper, H., Engler, H.: Mathematics and Climate. SIAM, Philadelphia (2013)
29. Kärner, O.: On nonstationarity and antipersistency in global temperature series. J. Geophys. Res. **107**(D20), ACL-1 (2002)
30. Kolmogorov, A.N.: Wiensche Spiralen und einige andere interessante Kurven im Hilbertschen Raum. Comptes Rendus (Doklady) Acad. Sci. URSS **26**(2), 115–118 (1940)
31. Koutsoyiannis, D., Montanari, A.: Statistical analysis of hydroclimatic time series: uncertainty and insights. Water Res. Res. **43**(5), W05429 (2007)
32. Kuehn, C.: A mathematical framework for critical transitions: normal forms, variance and applications. J. Nonlinear Sci. **23**(3), 457–510 (2013)
33. Kuehn, C.: Multiple Time Scale Dynamics, Volume 191 of Applied Mathematical Sciences. Springer, Cham (2015)
34. Kuehn, C., Malavolta, G., Rasmussen, M.: Early warning signs for bifurcations with bounded noise. J. Math. Anal. Appl. **464**(1), 58–77 (2018)
35. Kuznetsov, Y.A.: Elements of Applied Bifurcation Theory, 3rd edn. Springer, New York (2004)
36. Maslowski, B., Nualart, D.: Evolution equations driven by a fractional Brownian motion. J. Funct. Anal. **202**(1), 277–305 (2003)
37. Mishura, Y.: Financial applications of fractional Brownian motion. Stochastic Calculus for Fractional Brownian Motion and Related Processes, pp. 301–326. Springer, Berlin (2008)
38. Moberg, A., Sonechkin, D.M., Holmgren, K., Datsenko, N.M., Karlén, W.: Highly variable Northern Hemisphere temperatures reconstructed from low- and high-resolution proxy data. Nature **433**(7026), 613 (2005)
39. Molchan, G.M.: On the maximum of a fractional Brownian motion. Theory Probab. Appl. **44**(1), 97–102 (2000)
40. Nourdin, I.: Selected Aspects of Fractional Brownian Motion. Springer, Mailand (2012)
41. Odde, D.J., Tanaka, E.M., Hawkins, S.S., Buettner, H.M.: Stochastic dynamics of the nerve growth cone and its microtubules during neurite outgrowth. Biotechnol. Bioeng. **50**(4), 452–461 (1996)
42. Palmer, T., Williams, P.: Stochastic Physics and Climate Modelling. Springer, Cham (2010)
43. Piterbarg, V.I.: Asymptotic Methods in the Theory of Gaussian Processes and Fields. American Mathematical Society, Providence (1996)
44. Rangarajan, G., Sant, D.A.: A climate predictability index and its applications. Geophys. Res. Lett. **24**(10), 1239–1242 (1997)
45. Richard, A., Orio, P., Tanré, E.: An integrate-and-fire model to generate spike trains with long-range dependence. J. Comput. Neurosci. **44**, 1–16 (2018)
46. Romano, F., Kuehn, C.: Analysis and predictability for tipping points with leading-order nonlinear terms. Int. J. Bifurc. Chaos **28**(8), 1850103 (2018)
47. Sadhu, S., Kuehn, C.: Stochastic mixed-mode oscillations in a three-species predator-prey model. Chaos **28**(3), 033606 (2017)
48. Slepian, D.: The one-side barrier problem for Gaussian noise. Bell System Tech. **41**(2), 436–501 (1962)
49. Sonechkin, D.M.: Climate dynamics as a nonlinear Brownian motion. Int. J. Bifurc. Chaos **8**(04), 799–803 (1998)

50. Stone, H.: Calibrating rough volatility models: a convolutional neural network approach. Quant. Financ. (2019) (to appear)
51. Su, J., Rubin, J., Terman, D.: Effects of noise on elliptic bursters. Nonlinearity **17**(1), 133–157 (2004)
52. Rootzén, H., Mikosch, T., Resnick, S., Stegeman, A.: Is network traffic appriximated by stable Lévy motion or fractional Brownian motion? Ann. Appl. Probab. **12**(1), 23–68 (2002)
53. Tikhonov, A.N.: Systems of differential equations containing small parameters in the derivatives. Mat. Sbornik N. S. **31**(3), 575–586 (1952)
54. Verhulst, F.: Methods and Applications of Singular Perturbations: Boundary Layers and Multiple Timescale Dynamics. Springer, Cham (2005)
55. Weiss, M.: Single-particle tracking data reveal anticorrelated fractional Brownian motion in crowded fluids. Phys. Rev. E **88**(1), 010101 (2013)
56. Wouters, J., Lucarini, V.: Disentangling multi-level systems: averaging, correlations and memory. J. Stat. Mech. **2012**(03), P03003 (2012)
57. Xu, Y., Guo, R., Xu, W.: A limit theorem for the solutions of slow-fast systems with fractional Brownian motion. Theor. Appl. Mech. Lett. **4**(1), 013003 (2014)
58. Yuan, N., Fu, Z., Liu, S.: Extracting climate memory using fractional integrated statistical model: a new perspective on climate prediction. Sci. Rep. **4**, 6577 (2014)